Taxonomy of Fundamental Concepts of Localization in Cyber-Physical and Sensor Networks

Anis Koubâa · Maissa Ben Jamâa

Abstract Localization is a fundamental task in Cyber-Physical Systems (CPS), where data is tightly coupled with the environment and the location *where* it is generated. The research literature on localization has reached a critical mass, and several surveys have also emerged. This review paper contributes on the state-of-the-art with the proposal of a new and holistic taxonomy of the *fundamental* concepts of localization in CPS, based on a comprehensive analysis of previous research works and surveys. The main objective is to pave the way towards a deep understanding of the main localization techniques, and unify their descriptions. Furthermore, this review paper provides a complete overview on the most relevant localization and geolocation techniques. Also, we present the most important metrics for measuring the accuracy of localization approaches, which is meant to be the gap between the real location and its estimate. Finally, we present open issues and research challenges pertaining to localization. We believe that this review paper will represent an important and complete reference of localization techniques in CPS for researchers and practitioners and will provide them with an added value as compared to previous surveys.

Keywords Fundamental techniques of localization · Localization accuracy metrics · Localization real-world challenges · Localization open issues

1 Introduction

Cyber-physical systems (CPS) have been attracting an increasing interest from both the academic and industrial community. Indeed, these systems are showing powerful potentials in interfacing with the physical-world and making its control much easier. This achievement was enabled by the integration of computation and communication capabilities to the components of this physical world. Cyber-physical systems are representing today the new generation of networks and embedded systems. The deployment of CPS witnesses several challenges, and it is not new to say that localization is one of the most important topics that has triggered huge amount of research. Basically, localization aims at determining the position of CPS components. This topic has been considered by the CONET consortium as the most important in cooperative object research [16]. The amount of works related to localization is so tremendous such that it becomes challenging to a (novice) researcher to have a global view of the different approaches that have been proposed. There has been some surveys [2,10,22,27,28,30,32,51,56,76] covering the localization problem but each of them tackles it for a different perspective. One may ask: yet, another survey on localization? When looking at these surveys, it becomes clear that there is a big overlap between them, however, some of them do not speak the same language, in the sense that there is no unified terminology/taxonomy that gathers all these works. This fact represents one of the major motivation behind writing this survey with the aim to provide a unified taxonomy of localization for CPS. Our survey differs from others in several fronts. First, unlike some topic-oriented surveys, such as [22,28,30,32,76], this paper presents a comprehensive and generic overview of localization techniques in CPS. Indeed, the authors in [30,76] focused on range-free localization techniques, whereas the authors in [28] focused on TOA-based localization techniques. In [22], the authors were interested in localization techniques for underwater acoustic networks, and [32] addressed map-based localization with a particular interest on fingerprinting methods namely deterministic, probabilistic and filtering approaches. On the other hand, this survey paper contributes to the state-of-the-art as compared to other generic and thorough surveys, such as [2,10,27,51,56,82], in the sense that it provides a comprehensive taxonomy of localization techniques and discusses in details the fundamental concepts of each techniques as well as its features and application contexts. Furthermore, the current survey paper overviews localization accuracy metrics, which were not exhaustively discussed in other papers, to the best of our knowledge. As such, we believe that our survey would provide a key reference in the literature of localization in CPS. To meet its objective, this survey (i) makes a comprehensive review of fundamental localization techniques based on a fine grained analysis of the literature, (ii) designs a taxonomy in a structured way, (iii) elaborates on the applicability of these techniques in the CPS context, (iv) elaborates on the most representative metrics quantifying the accuracy of localization systems, and (v) discusses the future directions in localization research. Providing such information will help researchers and localization systems designers to understand the scope of different techniques, and to design appropriate localization algorithms for their systems.

The remainder of this survey is organized as follows. Section 2 presents a global taxonomy of fundamental localization concepts in CPS. For each category we examine its features and application context and we enumerate their advantages and limitations. In Section 3, we present a thorough review of fundamental localization techniques, namely the three main classes: (i) range-based, (ii) range-free and (iii) geolocation techniques. Given the diversity of localization techniques and the need for evaluating and comparing their performance, we present in Section 4, the most relevant metrics that quantify their degree of accuracy. Finally, Section 5 wraps-up the main lessons of this survey paper, presents real-world challenges and describes the future research directions in this research area. In Table 1, we present the organization of this paper.

2 Taxonomy

Localization systems consist commonly of two main blocks: (i) the set of deployed nodes and (ii) the localization algorithm. Deployed nodes may have different *states*. A state refers to whether a node initially knows or manages to know its location during the execution of the localization algorithm. According to [10], there are basically three states: unknown, settled and beacon. At the startup of the localization algorithm, nodes can either be in state beacon or unknown. Beacon nodes, also referred to as landmarks or anchors, are those that already know their locations through a manual placement or through GPS reading. In contrast, unknown nodes, referred to as free, dumb, or target nodes, are those that do not have any information about their geographic locations. Over the time, unknown nodes may change their states to settled if they succeed to determine or estimate their locations. Both beacon and settled nodes are very useful for unknown nodes in order to estimate their locations as they could be considered as references. On the other side, localization algorithms aim at finding the location/position of unknown nodes. There are several ways to determine the location depending on the objective of the cyber-physical application and the underlying technologies. Global Positioning System (GPS) has been commonly assumed as the intuitive solution to determine accurately locations. Nevertheless, GPS is not always the most effective solution for CPS, due to cost and energy constraints. As alternatives, many other localization techniques have been proposed. These techniques localize unknown nodes by exploiting (i) the sensing and (ii) the wireless communication capabilities of CPS components. For example, some techniques use wave propagation characteristics such as the Received Signal Strength (RSS), or propagation delay (known as time-of-flight) to infer the distances/angles to some reference nodes and then estimate the location through simple geometric computations (e.g. lateration/triangulation techniques). Some other techniques estimate distances by exploiting the difference between heterogeneous waves' propagation properties (e.g the reception time difference of acoustic and radio waves (TDoA) [29,64,84]). Another class of localization algorithms, known as range-free localization [30,76], does not rely on distance estimation. In contrast, it determines unknown node position based on proximity/connectivity information or based on artificial generated events. Taking into account the diversity of localization schemes, it is utmost important to derive the key criteria enabling their categorization which facilitates, consequently, their design and the understanding of their pros and cons. To meet this requirement, the following taxonomy, depicted in Fig. 1, has been devised after the compilation of several research and survey papers [2,10,19,24,30,32,51,54,60,65,71,76].

2.1 Topology

The topology of a localization algorithm refers to where and how the location of a given node is calculated. In fact, locations can either be computed at node level or at a central unit level, and this is based on signal measurements either received from or reported by anchor nodes, respectively. The choice of the topology typically depends on the cyber-physical application, the computational capabilities, the nature and the configuration of the objects involved in the localization process. There are mainly four possible topologies [19,51].

– *Remote positioning*: The location of a node is computed at a central base station, where anchor nodes collect transmitted radio signals from the mobile node and forward them to the base station. The latter computes the estimated position of the mobile node based on the measured signals forwarded by the anchors.

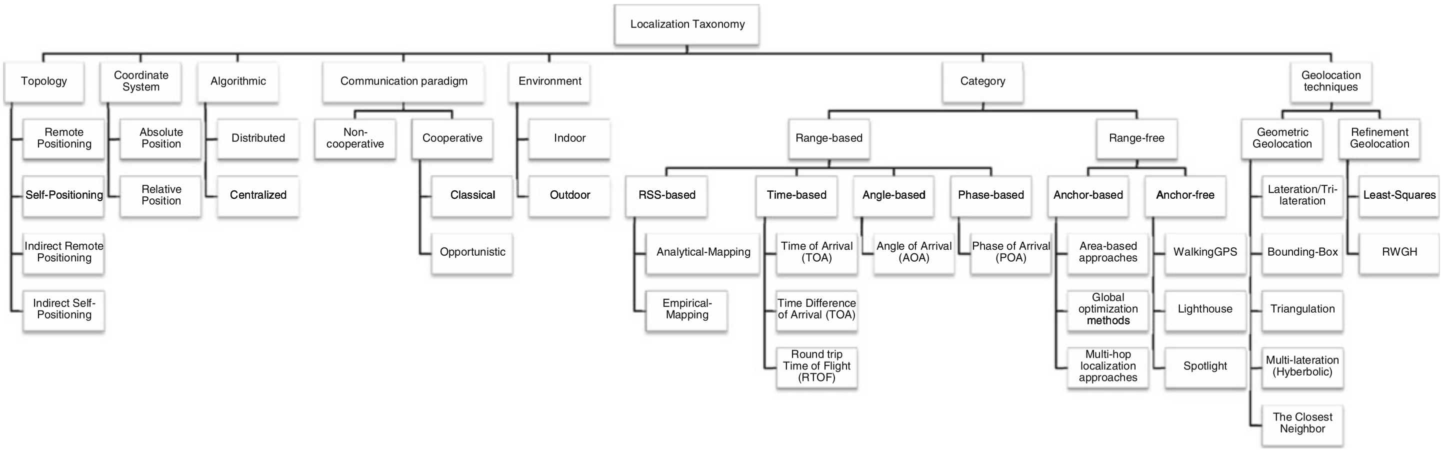


Fig. 1 Localization taxonomy

Table 1 Content of the paper

Topic	Section
Taxonomy	2
Topology	2.1
Coordinate system	2.2
Algorithms	2.3
Communication paradigm	2.4
Environment	2.5
Category	2.6
Fundamental techniques of localization	3
Range-based techniques	3.1
RSS-based localization	3.1.1
Time-based localization	3.1.2
Angle-based localization	3.1.3
Phase-based localization	3.1.4
Range-free techniques	3.2
Anchor-based methods	3.2.1
Anchor-free methods	3.2.2
Geolocation techniques	3.3
Geometric geolocation techniques	3.3.1
Refinement geolocation techniques	3.3.2
Localization accuracy metrics	4
Distance-based	4.1
Average relative deviation (ARD)	4.1.1
Global energy ratio (GER)	4.1.2
Frobenius (FROB)	4.1.3
Location-based	4.2
Mean absolute error (MAE)	4.2.1
Root mean square error (RMSE)	4.2.2
Cramer-Rao lower bound (CRLB)	4.2.3
Boundary alignment ratio (BAR)	4.2.4
Area-based	4.3
Circular error probability (CEP)	4.3.1
Geometric dilution of precision (GDOP)	4.3.2
Tile and distance accuracy	4.3.3
Discussions and future directions	5

– *Self-positioning*: The location of a mobile node is computed by itself based on the measured signals that it receives from anchor nodes.
– *Indirect remote positioning*: Like in the self-positioning case, the mobile node computes its locations and sends it to a base station through a wireless back channel.
– *Indirect self-positioning*: Like in the remote positioning, the location of the mobile node is computed at the base station based on measured signals forwarded by anchor nodes, then the base station forwards the estimated position to the mobile node, through a wireless back channel.

2.2 Coordinate System

The choice of the coordinate system type is fundamental in the design of localization algorithms. This choice is biased by the type of the information required to represent a location. The coordinate system can be either *absolute* or *relative*. For absolute coordinate systems, locations are expressed as unique coordinate values making reference to special nodes that know their positions, which are mainly the anchor nodes. At least, three anchors are needed in a 2D coordinate system, whereas four anchors are needed in a 3D coordinate system to be able to devise the locations of unknown nodes. For instance, GPS systems [61] use 24 to 32 satellites as anchor nodes to estimate the absolute location of GPS-enabled devices in outdoor environments. On the other hand, in relative coordinate systems, the location of a given node is determined relatively to other nodes with no reference to absolute anchors. In other words, reference nodes are not absolute in a relative coordinate system. For instance, Multi-Dimensional Scaling technique (MDS-MAP [73]) produces a map of relative coordinates based on mutual nodes' distance information [73].

Another classification of coordinate systems is related to the nature of coordinates, which can be either *physical* or *symbolic*. Physical locations are represented as a point in a 2D/3D coordinate system, such as the Universe Traverse Mercator (UTM) system, or the Degree/Minutes/Seconds (DMS) used for expressing GPS-based locations. On the other hand, symbolic locations are expressed as logical positions' information such as cell number, office/building number, or street name, etc.

2.3 Algorithms

Techniques of CPS objects location computation can be roughly classified into two categories: *distributed* and *centralized*. In centralized localization approaches, a central device (e.g. base station) is the responsible entity for estimating the location of unknown nodes based on the signal measurements forwarded by anchor nodes. In distributed localization approaches, each object estimates its location using the collected signal measurements and location information of the anchor nodes in its neighborhood. In the literature, distributed approaches are also referred to as *localized* algorithms [39]. With respect to topology, remote positioning and indirect self-positioning represent centralized algorithmic approaches, whereas self-positioning and indirect remote positioning represent distributed approaches.

For cyber-physical applications, centralized algorithms are appropriate for applications with a central monitoring station, which collects information from all objects in the network, such as surveillance systems, health care monitoring, environment monitoring, assets tracking. On the contrary, distributed algorithms are more convenient for decentralized applications such as swarm robotics, object auto-navigation, GPS-based systems.

Centralized and distributed approaches have opposite performance: centralized algorithms provide better localization accuracy at the cost of higher computation complexity and energy efficiency as compared to distributed algorithms [54,57]. In fact, centralized localization approaches are more accurate than distributed approaches because they rely on a global information, which is the sensory data collected from and forwarded by individual sensor nodes to the base station, to estimate the location. However, this centralization inherently induces a negative impact on scalability and computation efficiency. In contrast, distributed algorithms simply rely on local information to directly estimate the position rather than piggybacking data to a central unit. Therefore, the computation complexity of distributed localization algorithms is much simpler than that of centralized algorithms. The implementation of distributed algorithms in large-scale is also much easier. Regarding energy-efficiency, distributed algo-

Table 2 Comparison between distributed and centralized techniques

Performance criteria	Distributed	Centralized
Computation level	Node	A central unit
Measurements sources	Neighbors	All nodes
Accuracy	Medium	High
Scalability	High	Low
Energy consumption	Low	High
Communication overhead	Low	High
Computation overhead	Low	High
Computation delay	Low	High
Extra processing center	No	Yes
Anisotropic topologies	Robust	Sensitive

rithms are basically more energy efficient than centralized algorithms as the latter is subject to multi-hop communication between sensor nodes and the base station; whereas the former only requires one hop communication for location estimation. Nevertheless, the energy consumption in distributed algorithms may get higher if several iterations are needed to reach a stable estimation of the position [57]. A comparative performance study between distributed and centralized approaches has been presented in [67] and showed that the distributed algorithms are much more efficient, in terms of energy and communication than centralized estimation schemes using both analytical and simulation models. Table 2 summarizes the major differences between distributed and centralized localization approaches.

2.4 Communication Paradigm

When building a localization system, it is essential to define how nodes exchange messages between each other, which is referred to as *communication paradigm*. In the literature, there are mainly two communication approaches [82]: *non-cooperative* and *cooperative*. In the non-cooperative approach, the communication is restricted between unknown nodes and anchors, and there is no communication between nodes with unknown locations. In this case, a high density of anchors or long-range anchor transmissions [82] are needed to ensure that each unknown node is within the communication range of at least three anchors. On the other hand, the cooperative communication approach allows communication between unknown nodes in addition to communication between anchors and unknown nodes. As such, the need for high anchor density is alleviated. However, one major problem with cooperative localization is the need for intensive processing operations in order to filter noisy measurements collected at intermediate unknown nodes communication stage.

Another variant of cooperative localization is called *opportunistic localization* [88]. Unlike classical localization algorithms that assume homogeneous nodes and the deployment of a dedicated infrastructure, opportunistic localization exploits interactions between existing nodes and other nodes (which may be of heterogeneous nature) that occasionally pass in their proximity. Opportunistic localization raises a number of research challenges, mainly the efficient discovery of occasional nodes, the establishment of links between heterogeneous devices for opportunistic data exchange, and more important, the design of suitable protocols for efficient data exchange [88].

2.5 Environment

The environment plays an important role in the design of localization algorithms depending on whether it is *outdoor* or *indoor* environment. In outdoor environments, the radio propagation fits much better the free space propagation model than it does in indoor environments, due to the absence of obstacles and the little impact of interferences and multi-path propagations. For that reason, localization algorithms based on Radio Frequency (RF) signals are more convenient for outdoor environments. For instance, GPS-based systems are exclusively used for outdoor localization as GPS signals are not able to resist to obstacles. For indoor environments, the path loss propagation model does not hold and the RF-based localization becomes more challenging due to external factors of signal distortion mainly resulting from multi-path propagation. However, for low-power radios, the multi-path effect can be reduced by controlling the transmission power as the increase of the transmission power may result in higher-intensity destructive signals at the receiver. Interferences and noise impose additional challenges for indoor RF-based localization.

Other techniques for indoor localization rely on the Time Difference of Arrival (TDoA) mechanism, such as Cricket indoor location system [66], which combines RF and ultrasound signals to estimate a location. The key idea exploits the difference in propagation time between RF signal (speed-of-light) and ultrasound signal (speed-of-sound), and it has been shown that it provides very accurate location with an error around a few centimeters [57].

There are also other techniques that bypass the computation of distance and use neighborhood information to estimate the position of a node, known as range free techniques (refer to Sect. 3.2 for more details). These localization mechanisms are suitable for indoor environments as they are less complex and the location error would be tolerable taking into account the small scale and short radio ranges in such environments.

2.6 Category

To design a localization scheme for a given cyber-physical application, it is important to decide the category of the localization technique, which pertains to how the location of a node is calculated depending on whether they are based on distance measurement or not. In this respect, there are two main categories of localization techniques: *range-based* and *range-free*.

Range-based (or distance-based) techniques rely on the computation of distances between the target node and anchor nodes to infer the position of a target node using lateration techniques. Basically, distance measurement is achieved through RF signals or ultrasound signals or a combination of both. These techniques exploit the intrinsic propagation signal properties at the receiver to infer the relative distance of the target node through an empirical or analytical relationship/mapping between the received signal and the relative distance. Therefore, range-based localization requires complex computations to achieve high accuracy, which is considered as challenging for resource-constrained cyber-physical devices (e.g. sensor nodes) making hard to envision them as practical solutions for large-scale networks, in particular in noisy environments.

As another alternative, range-free solutions have been proposed. These approaches estimate the location of a target node without need to calculate distances but rather relying on other logical information including radio connectivity, anchor proximity and sensing capabilities (e.g. event detection). In [76], range-free techniques are classified as *anchor-based* schemes, assuming the existence of nodes with known positions, and *anchor-free* schemes that do not rely on any anchor node for localization.

Range-based solutions are known to achieve high localization accuracy at the cost of increasing system complexity in terms of ranging hardware, careful calibration and environment profiling. Being more tolerant in terms of accuracy, range-free solutions is likely to be more convenient for a large-scale networks of low-cost nodes as they do not require high cost specialized hardware for localization. Details about those techniques are presented in Section 3.

3 Fundamental Techniques of Localization

In this section, we describe the fundamental techniques used for calculating the location of a node according to the taxonomy presented in Fig. 1. First, we present the main approaches used for range-based and range-free schemes.

3.1 Range-Based Techniques

Range-based techniques for distance estimation can be categorized into four classes: (1) RSS-based, (2) Time-based, (3) Angulation-based and (4) Phase-based techniques.

3.1.1 RSS-Based Localization

These techniques are based on creating a mapping between the distance and the received signal strength as the signal attenuates when the distance increases. Ideally, for a certain environment, the distance to RSS mapping can be represented by the path-loss propagation model where the RSS is inversely proportional to d^{η}, where d refers to the distance between the transceivers and η refers to the path loss exponent, which is an environment-dependent parameter that represents the intensity of the signal attenuation in a given environment. Table 3 outlines a set of path loss exponents values for certain wireless networks. Equation (1) presents the general path-loss model expression:

$$RSS = P_t \times K \times \left(\frac{d_0}{d}\right)^{\eta} \tag{1}$$

where P_t is the transmission power, K is a constant that depends on the sender/receiver antenna gains, the wavelength and the path loss up to a reference distance d_0. Unfortunately, RSS is inherently unreliable as it gets affected by the random multi-path effect due to several physical phenomena of the signal propagation including reflection, refraction, diffraction and scattering. These phenomena results from the obstruction of physical objects during the signal propagation, also known as the *shadowing* effect making the signal weaker and more exposed to errors. RSS/Distance mapping methods can be roughly classified into *Analytical-Mapping* and *Empirical-Mapping* methods. A comparison between these two methods is presented in Table 4.

Analytical-Mapping models map RSS to distance through a mathematical equation. These models are commonly used in simulation software to emulate the channel behavior. The most common RSS to distance mapping model is the *log-normal distance path loss* propagation model, which is expressed as follows:

$$RSS(d)[dB] = RSS(d_0) - 10\eta \log(d/d0) + X_{\sigma}[dB] \tag{2}$$

where $RSS(d)$ is the received signal strength at distance d from the sender, $RSS(d_0)$ is the received signal strength at a reference distance d_0 from the sender fixed and known in

Table 3 Path loss exponent values for different types of environments [15]

Environment	Path loss exponent
Free space	2
Urban area (cellular radio)	2.7–3.5
Shadowed urban area (cellular radio)	3–5
In-building LOS	1.6–1.8
Obstructed in-building	4–6

Table 4 Comparison between RSS-based techniques

Technique	Advantages	Limitations
Analytical-mapping models	Simple to implement Useful for simulators design	Parameters are environment-dependent Coarse accuracy
Empirical-mapping models	Can achieve high accuracy level	Need extensive environment profiling High off-line computation overhead Poor scalability Unreliable if the environment is continually changing

advance, η represents the path loss exponent that measures the rate at which the received signal strength decreases with distance and X_σ is a zero-mean Gaussian random variable with a variance σ^2, which is referred to as the shadowing variance. Both η and σ^2 are environment dependent. Although the above equation seems to give a reasonable relation between the RSS and the distance, in reality, the establishment of the mapping is challenging and complex. We report three typical empirical observations to describe the challenging problem of finding a relation between the distance and the RSS.

Observation 1. The distribution of the RSS is not necessarily Gaussian: The path-loss log normal shadowing model assumes that the uncertainty is modeled as a white noise through the variable X_σ. This assumption does not hold in all situations, as the empirical distribution of the RSS shows up to be different in several cases. Indeed, this distribution typically depends on the environment and on the interference factors affecting the signal propagation and may not follow the normal distribution. Figure 2 illustrates our statement.

The figure shows the distribution of the RSS in two environments (indoor and outdoor) and with two transmission powers (0 and -15 dBm). We observe that probability distribution of the RSS depends on the environment and on the transmission power. Furthermore, it is clear from the figures that the RSS distributions do not match well the Gaussian distribution which compromises the validity of the assumption of the path loss log normal shadowing model (i.e. X_σ is a zero-mean Gaussian random variable). The reasons behind this discrepancy are manifold: (i) sensor node hardware imperfections or (ii) dynamic signal distortion caused random multi-path propagation, or by the moving objects/persons around sensor nodes or due to weather conditions. We can also observe that the RSS distribution is closer to Gaussian in outdoor environment than in indoor one. This argues more the effect of multi-path propagation caused by walls in the indoor environment.

Observation 2. The RSS variability is typically (very) high: On the other hand, in the path loss shadowing model, the variability of the RSS is modeled through the variance of the Gaussian noise. It is commonly known that the shadowing at a given fixed distance variance

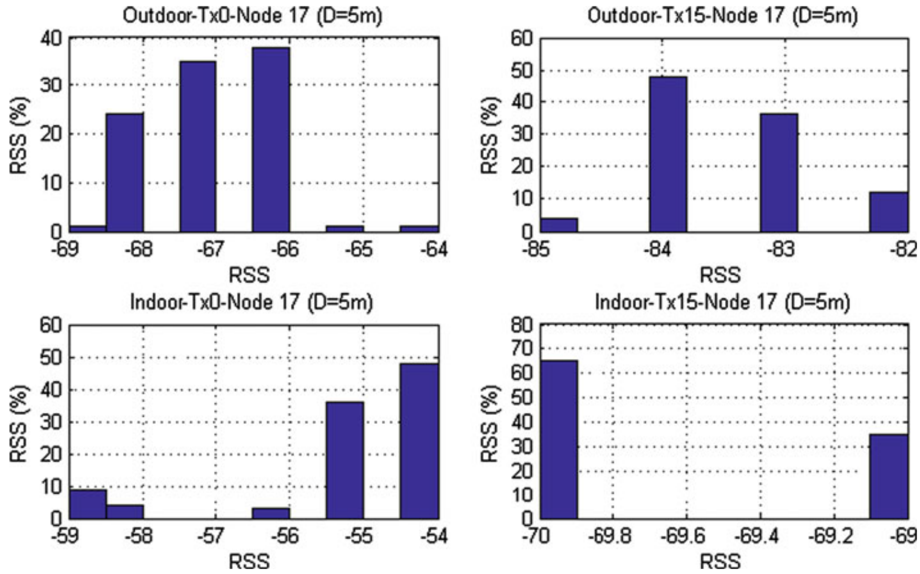


Fig. 2 RSS distribution at 5 m in different environments with Tx power = 0 and −15 dBm

might be high, in particular when the propagation is disturbed by the multi-path effect and interferences. Figure 3 demonstrates this fact and shows the RSS variance for one indoor and one outdoor environments. We observe that the RSS standard deviation in the indoor environment (with an average 3.78) is a bit higher than that in the outdoor environment (with an average 2.85) for the same reasons mentioned above. In addition, the non regularity of the radiation pattern of radio signals represents a main source of RSS variance in the same distance, which we discuss in the next observation.

Observation 3. The non isotropic behavior of signal propagation is the cause of RSS variability: One of the drawbacks of the path loss model is that it does not consider the relative orientation of the sender and receiver. The anisotropic behavior of the RSS turns the mapping with distance even harder, as radio propagation is not omni-directional, and thus the RSS may be highly variable at receivers placed at the same distance, but in different orientations. Figure 4 illustrates the problem and shows the RSS at different orientations and at the same distance.

In Fig. 4, we observe that the non-uniform radiation pattern of the sensor node transceiver makes the extents of the RSS variation interval different from one orientation to another. This is one the factors that comprises the localization accuracy based on the RSS.

It is clear from the aforementioned observations that log-normal distance path loss propagation model expressed in Eq. 2 does not reflect the reality in strict sense, but rather provides an acceptable approximation. Nonetheless, there have been several research works that used this model for localization purposes. Since η and σ^2 are environment dependent, usually it is needed to determine these parameters prior to deployment using extensive experimental tests, which may restrain the use of this model in reality. This practice is commonly referred as environment *profiling* or *fingerprinting* or *scene analysis*. However to avoid prior environment investigation, some research works have provided practical methods to dynamically adapt the theoretical log-normal propagation model to the environment of deployment. This adaptation focuses especially on the empirical estimation of the path-loss exponent. For

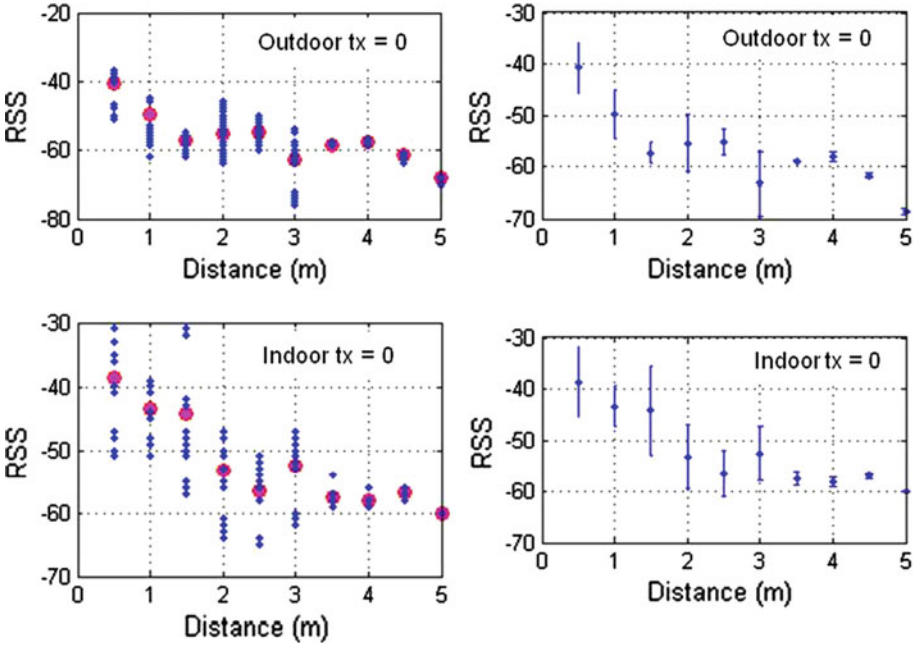


Fig. 3 RSS average and its standard deviation as a function of distance

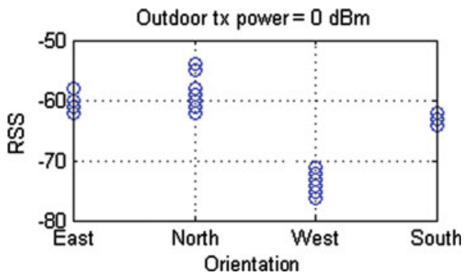


Fig. 4 RSS measurements for the same distance (3 m) at different orientations

instance in [77], the path-loss exponent has been estimated, based on the RSS measurements collected during the localization process and by assuming the existence of N fixed anchors evenly distributed. The drawback of this solution is that it relies on an approximated distance function, which induces additional complexity and less accuracy. In [55], the calibration of the path loss exponent has been conducted through two techniques. In the first technique, the path loss exponent is estimated using a small number of received power measurements and by assuming that the probability distribution of distance between neighboring sensors is known. However, this assumption is unrealistic for certain applications. To overcome this limitation, a second technique based on the Cayley-Menger determinant has been proposed. This technique estimates the path loss exponent using only power measurements and the geometric constraints associated with planarity in a cyber-physical system. In [5] too, the costly off-line fingerprinting procedure has been avoided and replaced by a virtual calibration procedure which exploits only RSS measurements. In this calibration procedure, the authors have used an enhanced propagation model which takes into account the wall and floor attenuation factors for indoor environments. Two estimations methods have been derived for single wall and multi-wall indoor environment models. The experimental study showed that the estimation accuracy was comparable to that achievable by a more computationally expensive fingerprinting procedure. Furthermore, a simpler yet effective approach has been proposed in [47]. This approach also exploits the theoretical propagation model for indoor localization with no manual profiling. The key idea makes use of Apollonius circles principle and the fact that the position of a node can be determined if the ratios of the distances to a few anchors are known. An Apollonius circle regroups all points having the same ratio of their distances to two fixed points. An unknown node is located at the intersection point of all circles for which it satisfies the Apollonius condition. In [6] too, a practical plug-and-play and distributed RSS-based localization method called EasyLoc was proposed. The idea of EasyLoc consists in exploiting the available distance information between anchors to derive an online and anchor-specific RSS-to-distance mapping. The main advantage of EasyLoc is its easy of deployment since it avoids any pre-deployment calibration phase and builds its mapping in runtime. This mapping is dynamically updated in order to be robust against environment change. Nonetheless, the mapping model used in EasyLoc is very simplistic and does not hold the real variation of RSS with respect to distance.

It appears from the aforementioned works that online and off-line calibrations of the theoretical path-loss propagation model are challenging, and the achievement of an optimized trade-off between accuracy and simplicity is a complex problem. However, for accuracy-tolerant cyber-physical applications, this propagation model would be sufficient to reach an acceptable accuracy with simple adaptation techniques. Alternatively to analytical-mapping models, empirical-mapping models, which we describe next, can also be used for RSS/distance mapping. These models are more accurate, but also more labor-intensive.

Empirical-Mapping models map RSS to distance through experimental measurements and statistical analysis of collected data. The most common technique is the *map-based* localization, which is mainly based on *fingerprinting* the environment through extensive pre-deployment measurements [32]. This technique is composed of two phases as depicted in Fig. 5. (i) the *training phase*: it consists in measuring the RSS at different locations in the deployment area, then forming a *radio map*, which represents the mapping between locations and their corresponding measured RSS, (ii) the *positioning phase*: A node with unknown location will be able to localize itself by comparing its RSS with those in the map, and then estimates its position as being the location corresponding to the closest RSS in the map. The map-based technique presents the advantage of providing accurate results as compared with other techniques (e.g. analytical models), and this accuracy pertains to the amount of measurements collected during the training phase. However, it has the following drawbacks: first, the profiling operation is too complex, which increases with the area of the explored environment and the number of devices to be deployed in the training phase. Second, any change in the environment, such as the movement of persons/objects, will compromise the validity of the static mapping already established in the training phase, or resulting in increasing loss of accuracy. One way to reduce the human labor for environment fingerprinting is to automate the mapping process, which allows to perform frequent updates of the radio map. This technique is used in [83] for a cooperative robot/RFID system. Initially, the robot gets a radio map from a location server, which enables it to estimate its distance at a meter-level. Then, the robots move towards RFID tags, which serve as landmark as their location is known, and calculates its exact location. Using the exact location information, the robot can automate the training phase and reconstruct a new radio map to improve the precision of the fingerprinting algorithm.

The construction of the map starts with subdividing the area into regular cells. The collection process starts when Access Points transmit radio signals, which will be received by the

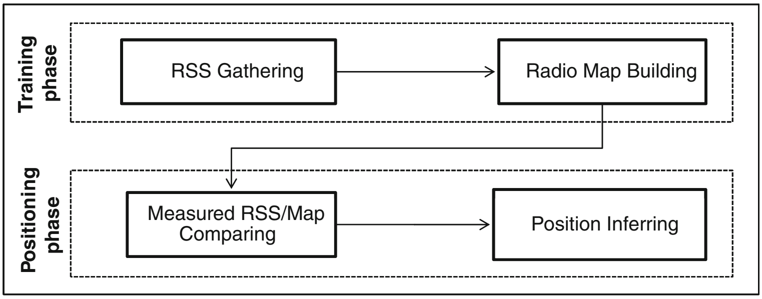


Fig. 5 Phases of fingerprinting

calibration devices for a certain period of time at a fixed location. The process is repeated at different locations in the area of interest. The received RSS will be stored in a vector data $\mathbf{a}_{ij}$ in a database, where i denotes the ith cell and j the jth access point. The collection of RSS vectors is called the *radio map*. The radio map can also include other information that would be useful for the localization process. For instance, in [1] the authors have considered three parameters in radio map including the central position of the ith cell, a vector $\mathbf{a}_i$ whose jth element represents the average RSS measured in cell i from Access Point j, and a diagonal matrix whose jth element represents the variance of the RSS values measured from the jth Access Point.

The most challenging issue in *map-based* localization is how to measure the similarity between the fingerprints and collected measurements [62]. The trivial approach consists in computing the minimum Euclidean distance between the observed RSS and the mean of each fingerprint such as in [37]. For instance, RADAR localization system [4] relies on the measurement of the Euclidean distance in order to find the k-nearest-neighbors to the target node and thus to be able to compute its location. Other sophisticated methods are recently proposed to better take into account the variability of the RSS. These methods generally rely on the use of the probability theory by first generating probability densities for the training data and then by computing the Maximum-Likelihood. Kernel-based nonlinear methods have also been investigated for similarity computation, such as in [41]. Nonetheless, these methods often require the collection of large data samples in the training phase and high processing capabilities [62], which is generally beyond the capacity of low-cost cyber-physical devices. Recently, a new methodology of exploiting fingerprints known as *Radio Tomographic Imaging (RTI)* has emerged [38,81]. In this approach, the fingerprint does not represent a vector of RSS values for a specific location, but it rather represents a vector of RSS values for a radio link in the network. To locate a target, an RTI system bases its detection of links that were attenuated. Radio Tomographic Imaging is still a challenging issue for indoor environments as the target presence is not the only factor affecting the RSS.

3.1.2 Time-Based Localization

It represents another class of radio-based distance measurement techniques, which rely on radio signal propagation time. In other words, time-based localization consists in calculating the distance by measuring the radio signal propagation time from the source to the destination. In the literature, there are, basically, three known kinds of time-based approaches. Figure 6 depicts them.

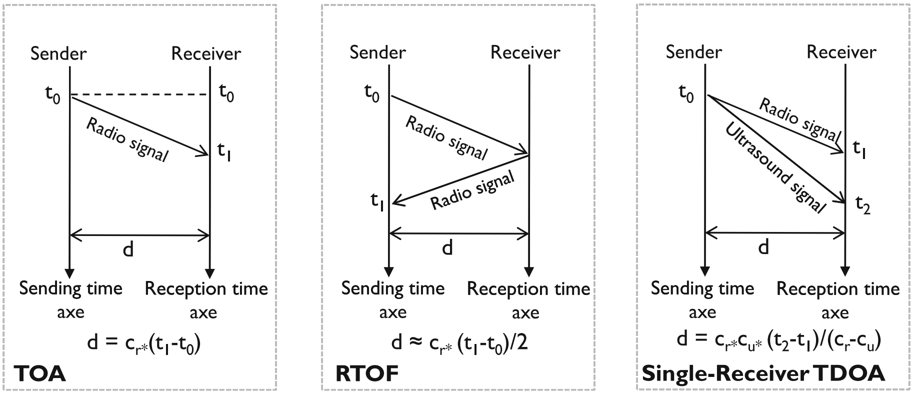


Fig. 6 Time-based localization techniques

*Time-of-Arrival (*TOA*)* Also called *Time-of-Flight* (*TOF*), TOA consists in calculating the one way propagation time of radio signals between two synchronized nodes. In fact, this time is proportional to the distance between transceivers since the propagation speeds of the RF signals are well-known both in free-space and in the air. This distance is simply given by Eq. 3 when no synchronization errors are available.

$$d = c_r \times (t_1 - t_0) \tag{3}$$

where c_r refers to the speed of the RF signal, t_0 to the transmission time and t_1 to the reception time.

The TOA method of RF signals is usually inappropriate for WSNs because of short distances and inaccurate time synchronization of sensor nodes. Acoustic or ultrasonic signals represents another alternative better than RF signals in order to use the TOA method in such systems. Nevertheless, RF-based TOA is typically applicable for GPS systems with large distances and high clock synchronization.

In general, the main disadvantages of this method for CPS are three folded: (i) transceivers of different nodes must be accurately synchronized, which requires high clock resolutions. (ii) the TOA accuracy depends on the RF bandwidth, which means that higher bandwidths provide better accuracy as it is the case with Ultra Wide Band Technology (UWB), (iii) they are very sensitive to multi-path effect as TOA represents a direct line-of-sight propagation time, thus the blockage of the direct path will cause large errors. Some research works have proposed solutions to mitigate the non line-of-sight propagation problem [80].

The TOA method can be used with either Direct Sequence Spread Spectrum (DSSS) [58] or UWB radios [51]. With UWB technology, high accuracy can be achieved mainly because (i) there is more tolerance with respect to clock resolution as the propagation speed of ultrasound radios is relatively small (approximatively 331.4 ms/s [51]) (ii) it has a large bandwidth ($\geq$ 500 MHz). This turns the measurements of the transmission delay more accurate and inexpensive. It has been shown that the location estimation accuracy with UWB radios can be up to 2 cm in good conditions with direct line-of-sight propagation. With other RF-based TOA, the accuracy is roughly from 5 to 10 m.

*Round-trip Time of Flight (*RTOF*)* This is another variation of TOA, which attempts to avoid synchronization constraints. The main idea is to measure the round-trip TOA at the sender side, and since the same clock will be used for calculating the delay, the synchronization problem does no longer hold. However, the signal processing delay at the receiver side must be estimated and eliminated to prevent additional source of errors. The signal processing delay in the receiver can be pre-calibrated in advanced or measured by the receiver, and then sent back to the sender to be subtracted. Approximately, distance between the sender and the receiver can be given by Eq. (4):

$$d = c_r \times \frac{(t_1 - t_0)}{2} \tag{4}$$

where c_r refers to the speed of the RF signal and $(t_1 - t_0)$ to the round-trip time of flight.

It has to be noted that the accuracy of RTOF is also compromised by noise, multi-path effect and the unavailability of NLOS path.

*Time Difference of Arrival (*TDOA*)* Two different definitions of TDOA are available in the literature. The early approach used in geolocation systems, such as telecommunication and satellite communication systems [31], consists in calculating the time difference of arrival in two different receivers placed at different distances from the sender. We refer to this method as *multiple-receiver TDOA*. The other approach, being used in low power device systems, is based on the time difference resulting from the difference of propagation durations of two signals received by the same receiver. We refer to this method as *single-receiver TDOA*. In what follows, we present the concepts of both approaches and discuss their advantages and drawbacks.

The *single-receiver TDOA*: This approach is based on the measurement of the time difference of the propagation of two signals with very different propagation speeds. Doing so, the synchronization requirement of TOA method is bypassed. The most common signals used in TDOA-based localization systems are radio signals and ultrasound signals, as RF signals (speed-of-light) are roughly 10^6 faster than ultrasound signals. As a matter of fact, this combination is used in the commercially available sensor platform Cricket [64] and has been shown to provide an accuracy of 5 cm. Acoustic signals are also used as another alternative to ultrasound signals. This TDOA approach works as follows: When a node simultaneously sends an RF signal and an ultrasound signal, the receiver considers the arrival time of the (faster) radio signal as a time reference and uses the arrival time of the (slower) ultrasound radio to calculate the delay between both signals. The distance d between the sender and the receiver is calculated according to the following equation:

$$d = \frac{c_r \times c_u \times (t_2 - t_1)}{c_r - c_u} \tag{5}$$

where c_r and c_u are respectively the propagation speed of both radio and ultrasound signals, while t_1 and t_2 are their reception times at the receiver level. The TDOA method has the advantage to provide much better accuracy than RSS-based methods and it does not suffer from the need of explicit synchronization between nodes. However, it presents the drawback of requiring additional and more complex hardware with two different transceivers, which would have a negative impact on the cost.

The *multiple-receiver TDOA*: In this approach, the TDOA method is seen from another perspective. It is based on the propagation of only one signal (e.g. radio) and needs at least two receivers to estimate the reception time difference in the multiple receivers [51]. When a sender node transmits a signal, it will be received at different instants at two receivers located in different positions. The TDOA results from the distance difference of both receivers from the transmitter. Thus, it has been formally shown that for two receivers at known locations and with a known TDOA, the sender node is then located on a hyperboloid whose equation is given by:

$$R_{i,j} = D_i - D_j \tag{6}$$

where D_i is the distance between receiver i and the sender node defined as:

$$D_i = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} \tag{7}$$

where (x_i, y_i, z_i) represents the 3D coordinates of a fixed receiver i, and (x, y, z) represents the coordinates of the sender. Finding the exact solution of hyperbolas intersection is very complex. There is though linear techniques using Tailor-series to find an estimated location [78]. Another conventional method relies on computing the cross-correlation function of the signals arriving to two receivers and estimate TDOA being the time that maximizes the cross-correlation function [51].

3.1.3 Angle-Based Localization

This class mainly represents the *Angle-of-Arrival* (*AOA*) method also known as *Direction-of-Arrival* (*DOA*) or *bearing* measurements [57]. This method relies on computing the angle (or the direction) of the line connecting an unknown node to an anchor node with respect to some reference direction. The reference direction is also referred to as *orientation* [71]. If there is no reference direction (the orientation is unknown), the angle is defined by two lines connecting the unknown node with two anchor nodes.

The orientation is considered as *absolute* if it refers to the North direction, which means an angle of 0°. It is considered as *relative*, otherwise, where a relative reference direction with respect to the North is known in advanced. The concept of absolute and relative orientations is illustrated in Fig. 7. In the literature, it is assumed that each node may have its own orientation axis different from others' nodes orientations [60].

If the orientation of an unknown node is known (either absolute or relative), only two anchor nodes would be sufficient to estimate the position of the target. However, when the orientation is unknown, at least three non-collinear anchor nodes are needed for locating the unknown node, as illustrated in Fig. 7c. In this case, since the absolute angle information cannot be determined, the angle difference between two different nodes viewed by the third one is utilized instead.

In [87], the authors have proposed localization algorithms using AOA information in the case of unknown orientations. The first method relies on dual information pertaining to distance and AOA whereas the second method only considers AOA information. It has been shown that the knowledge of distance improves the robustness of the localization against noise but is inherently dependent on distance measurements.

Geometrically, the estimated angle $\hat{\theta}$ between a target node with coordinates (X_{target}, Y_{target}) and an anchor node with known coordinates (X_{anchor}, Y_{anchor}) is determined by [18,57]:

$$\hat{\theta} = \theta + n_i \tag{8}$$

where

$$\theta = \arctan\left(\frac{X_{target} - X_{anchor}}{Y_{target} - Y_{anchor}}\right) \tag{9}$$

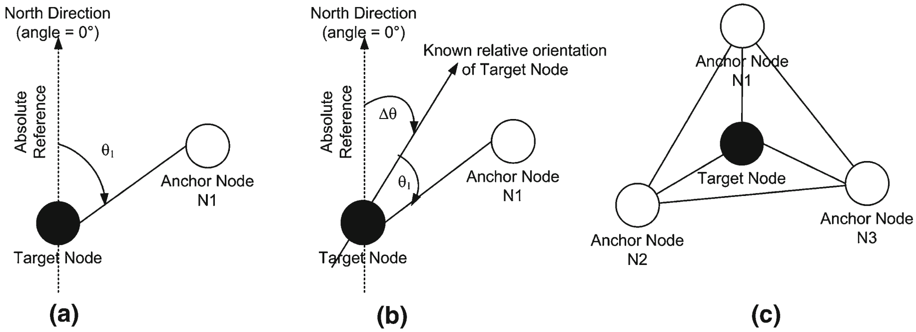


Fig. 7 AOA orientation concept. **a** Absolute orientation. **b** Relative orientation. **c** AOA with unknown orientation

where n_i is the additive zero-mean Gaussian noise with variance σ^2, and $\tan^{-1}$ is the inverse of trigonometry tangent function.

In practice, there are two fundamental techniques for measuring angles [39]. The first technique relies on the use of directional antennas that can rotate on their axis. While this technique makes simple the computation of angles, it is not practical for low-cost cyber-physical devices (e.g. sensor nodes) as their propagation pattern is typically assumed to be quasi omnidirectional. The second technique is based on the use of an array of antenna, which exploit the finite propagation speed of waves. Figure 8 illustrates this concept and shows an array of N antenna where adjacent ones are mutually separated by a distance d. The distance of the transmitter to the kth antenna is approximated by the following equation [57]:

$$R_k \approx R_0 - kd\cos(\theta) \tag{10}$$

where R_0 represents the distance of the transmitter to the 0th antenna, and θ is the direction of the transmitted signal viewed from the antenna array. The phase of the received signal is equal to $2\pi\frac{d\cos(\theta)}{\lambda}$ with λ is the wavelength of the transmitted signal. Thus, the AOA can be derived from the phase differences of received signals.

The advantage of the AOA method is that it does not require time synchronization in contrast to time-based localization methods (e.g. TOA). However, it exhibits several disadvantages making it not very appropriate for CPS. First, AOA requires very complex and expensive hardware which contrasts the basic requirements of low-cost, low-power devices. Second, the computation paradigm of angles is inherently complex, challenging and pretty much affected by noise, shadowing and multi-path reflections incident from misleading direc-tions. In the literature, it is very difficult to find research works on AOA using experimental validation of the proposed mechanisms as the majority of these works such as [18,60,71,87] rely on simulations for validation purposes.

3.1.4 Phase-Based Localization

This class mainly represents the Phase-of-Arrival (POA) method which is also called Received Signal Phase Method [51]. This method consists in finding the carrier phase of the received signal by the receiver and devise the corresponding distance. This method is effective only when the signal's wavelength is longer than the maximum distance to be estimated. In this technique, it is assumed that all nodes send sinusoidal signals of the form $S_i(t) = \sin(2\Pi ft + \varphi_i)$. The propagation delay is proportional to the phase

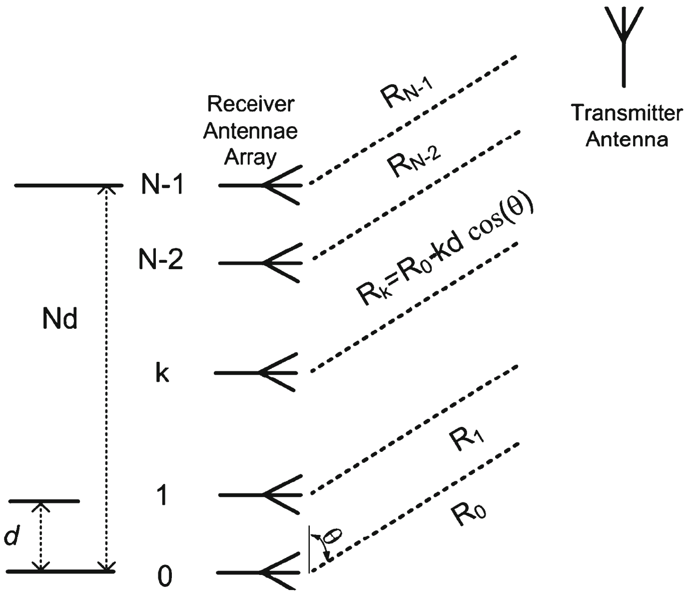


Fig. 8 Illustration of AOA concept of an array of antenna

$\varphi_i = (2\Pi f D_i)/C$, where C is the speed-of-light, and D_i is the delay between sender and receiver i. If the wavelength is longer than the maximum distance, then $0 \leq \varphi_i \leq 2\Pi$ and the delay can be estimated as $D_i = (C\varphi_i)/((2\Pi f)$. The POA method, like previous methods, needs the LOS path, otherwise errors will degrade its accuracy. In addition, carrier phase estimation is a challenging task making the practical use of this POA tricky. The POA method can be used in combination with the aforementioned methods to improve the accuracy of the localization. Table 5 presents a comparision between these range-based techniques.

3.2 Range-Free Techniques

Range-free methods do not rely on distance or angle estimation in localization. They rather use proximity or connectivity information to devise the location of the target. In [57], a classification of range-free techniques has been proposed and considers three classes: (i) Connectivity-based localization, (ii) non-parametric RSS-based localization, (iii) RF fingerprint-based localization. In the comprehensive survey about range-free localization methods [76], these methods have been classified into *anchor-based* and *anchor-free*, which we believe is more convenient for a proper taxonomy. In fact, RSS-based method and RF fingerprinting are typically used in range-based techniques, as described above, although distance can be ignored in some cases to estimate the location based on connectivity. However, this does not represent the essence of these techniques. In what follows, we present the most representative techniques of range-free methods according to the classification proposed in [76].

3.2.1 Anchor-Based Methods

These methods assume the existence of anchor nodes that know their locations. In a 2D space, three anchors are required, whereas in a 3D space, at least four anchors are needed to estimate a location of a target based on radio connectivity. Several techniques have been proposed in

Table 5 Comparison between range-based techniques

Localization methods	Advantages	Limitations
Time-based		
Time-of-arrival	Accurate	Node synchronization is mandatory
	Low communication and computation overhead	Require extra sensing hardware
Time-difference-of-arrival	Accurate Low communication and computation overhead Synchronization is not mandatory	Extra ultrasound transmitter is required
Round-trip time-of-flight	Accurate	Prune to NLOS effects
	Low communication and computation overhead Synchronization is not mandatory	
RSS-based	No need for extra sensing hardware	High variation especially in indoor environments
	Simple to implement Low cost	Sensitive to interferences and multi-path effects Unreliable and inaccurate
Angle-based	Give information on orientation Synchronization is not mandatory	Affected by NLOS and multi-path effects Sophisticated or arrays of antenna is required
Phase-based	Simple to implement	The signal's wavelength must be longer than the maximum distance to be estimated

the literature to approximate the location of a target based on connectivity information and without recourse to distance. In this paragraph, we present several anchor-based range-free localization techniques, which in turn can be classified as (i) *Area-based approaches* such as centroid and point-in-triangle techniques, (ii) *Global optimization approaches* such as Multidimensional Scaling (MDS), or (iii). *Multi-hop localization approaches*.

(a) *Area-based approaches*: these techniques are based on radio connectivity between the target and anchors and typically estimate the position of the target as a particular point in the polygon formed by all neighbors. Two main methods are proposed in the literature: (i) the centroid method, (ii) the point-in-triangle.

   (i) *Centroid*: this is an *area-based* technique that was first proposed in [11] and it merely consists in computing the location of the target as the *centroid* point, defined as the geometric center (or barycenter) of a set of anchors, as shown in Fig. 9.

In the most general form, assuming n anchor nodes A_i with coordinates (X_i, Y_i) are detected by the target node (through beacon listening), the latter calculates its coordinate (X_G, Y_G) such that:

$$(X_G, Y_G) = \left(\frac{\sum_{i=1}^{n}(X_i)}{n}, \frac{\sum_{i=1}^{n}(Y_i)}{n} \right) \tag{11}$$

To refine the localization accuracy using the centroid technique, a point mass (or a weight) can be adequately assigned to each anchor node depending on a predefined criterion. In this case, the coordinate of the target node is given by:

$$(X_G, Y_G) = \left(\frac{\sum_{i=1}^{n}(m_i X_i)}{\sum_{i=1}^{n}(m_i)}, \frac{\sum_{i=1}^{n}(m_i Y_i)}{\sum_{i=1}^{n}(m_i)} \right) \tag{12}$$

where m_i represents the point mass of the anchor node A_i, which defines the weight assigned to it. This technique is commonly known as the Weighted Centroid Localization (WCL).

In the literature, several studies have focused on finding optimal values for point masses based on the RSS or distance information [8,53,68] and several values were proposed. In [53], m_i was chosen to be equal to the inverse of the estimated distance, d_i, between the anchor node A_i and the target node, whereas reference [8] proposed more general weight having m_i equal to $\frac{1}{(d_i)^q}$, where q is a degree ensuring a greater impact of long distances of estimated target position. Similarly, in [68], the inverse of $(RSS)^q$ was used as weight for anchor nodes. In these papers, it has been shown that centroid localization provides better accuracy when used with adequate weights as compared to the general case, while still keeping low computation complexity.

(ii) *Approximate Point-in-Triangle Test* (*APIT*): It is another *area-based* localization technique proposed in [30]. It assumes that the location of the target is the center of gravity of a certain triangle, which is defined as the intersection of triangles formed by anchors in which the target node resides, as illustrated in Fig. 10. The idea simply consists in dividing the environment into triangular regions, then testing whether the target is inside a given triangle or not to narrow down the area of the possible target locations. All possible triangle combinations are tested: When the target is outside a given triangle, it will not be considered in the computation of location. At the end, the center of gravity of the triangles' intersection will represent the estimation target position.

The main challenge in this technique is to determine whether the target is inside or outside a certain triangle. In [30], the authors proposed an exact, yet theoretical, test that correctly decides whether a point is inside a triangle or not. The idea referred to as *perfect PIT* and relies on checking the existence of a direction such that if the target moves according to that direction, it will simultaneously get further/closer to all triangle points, i.e. anchors. This approach is illustrated in Fig. 11. Two major handicaps hinder the practicability of this approach in real-world. First, nodes typically do not have the ability to recognize the direction without moving. Second, it is not possible to perform an exhaustive test covering all possible directions in which the target may move to. To solve both problems, an approximation has been proposed and whose idea is to use neighborhood information, exchanged via beaconing, to emulate the node movement in the Perfect PIT test as shown in Fig. 12. First, the target node asks its neighbors for their distances to three *corner* anchors. The target then compares its distance to these three corner anchors against those of its neighbors. If there exists at least one anchor such that it is further from or closer to all corner anchors than the target, then the latter considers itself as being outside the triangle. Otherwise, if all neighbors are closer to some anchors and further from some others, then the target infers that it is located inside the triangle.

With respect to distance estimation and comparison, RSS has been proposed to be used. This produces errors in the APIT test leading to incorrect decisions. It has been reported in [30] that, based on experiments, the decision error of APIT does not exceed 14 % in the worst case. The APIT approach provides better accuracy in case of dense networks as the probability of such errors would decrease. For additional details about the performance of APIT and the comparison of its behavior against other localization techniques, the interested reader is referred to [30] and [76].

(b) *Global optimization methods*: Roughly, these methods assume the knowledge of global information to find an optimal solution to the main problem (e.g. inferred optimal location of nodes). *Multidimensional Scaling* (*MDS*) has been used as a global optimization technique for finding the location of nodes based on the global knowledge of distances between nodes. In a more general perspective, this technique is used in visual representation of the information for the sake of exploring similarities or dissimilarities in data. Applied to localization, it consists in inferring the different locations of nodes in multidimensional 2D or 3D space (visual representation of geographical data) based on the information of relative distances (global information input). In other words, MDS attempts to arrange nodes in a 2D or 3D space so as to reproduce a map such that the resulting distances on the map are as much close as possible to the observed distances. A typical example, is map reconstruction based on inter-city distances. Obviously, if the map containing multiple cities is known, it is straightforward to determine the relative distance between cities, which provides a unique solution. In contrast, if distances between cities are known, finding the exact location of each city in the map is too complex and not deterministic. MDS provides a solution to approximate the real positions based on distance observations.

Usually, MDS is formulated as an optimization problem, where the estimated locations of N nodes $(x_1, \ldots, x_N)$ are those that minimize some cost functions, such as $\min_{x_1,\ldots,x_N} \sum_{i<j} (\|x_i - x_j\| - d_{i,j})^2$, where $\|.\|$ is the vector norm operator and $d_{i,j}$ is the observed distance between node i and node j.

In [73], the authors proposed MDS-MAP, a method for using MDS in localization. MDS-MAP uses the Classical Multidimensional Scaling (CMDS), which is the simplest case of the MDS technique that does not require iterations as it provides a closed form solution. The network is assumed to be an undirected graph with vertices V and edges E. The vertices represent nodes whereas edges can either represent (i) connectivity information or (ii) distances between nodes, if those are known. In general, the MDS-based localization algorithm generates a relative map, which tailors a possible geographical visualization of the network based on distance or connectivity information. If, in addition, the absolute locations of some anchors are known, then the absolute map can also be derived. The proposed localization algorithm has three steps:

(i) *Shortest-path computation*: The algorithm first computes the shortest paths between all pair of nodes with either connectivity or distance information. The resulting distances in the shortest path are used to construct the distance Matrix.

(ii) *Relative mapping*: MDS is then applied to the distance Matrix and relative positions are computed. For m-dimensional space ($m = 2$ or 3) the coordinate matrix representing the location is based on the m eigenvalues and eigenvectors after the singular value decomposition. Although it is possible to achieve high accuracy in terms of resulting inter-node distance, the location of nodes will be arbitrary rotated and flipped as compared to real node locations.

(iii) *Absolute mapping*: Assuming that locations of some anchor nodes are known, the algorithm devises the absolute map through linear transformation including scaling, rotation, and reflection. The absolute locations are considered to be found when the sum of the squares of the errors between the true positions of the anchors and their observed location in the map is minimized.

More details about mathematical models of MDS-MAP can be found in [73,76]. It has been shown that the MDS-MAP localization has a complexity in the order of $\mathcal{O}(n^3)$ due to the singular value decomposition operation in MDS. This complexity cannot be handled by resource constrained nodes, thus, MDS computation must be executed on a base station with powerful computation capacity.

The main shortcoming of MDS-MAP method is the need for global information about inter-node distances or connectivity information. This is very improbable in case of dense wireless sensor networks due to energy and bandwidth constraints. In addition, MDS performs worse than other techniques in dense sensor networks. The knowledge of connectivity information is usually more feasible than the knowledge of real distances. However, this actually compromises the accuracy of relative and absolute computed geographical locations. Furthermore, this approach is centralized and is not convenient for mobile CPS, in general, as distributed approaches would be more suitable.

In the literature, several techniques have been proposed to enhance the performance of MDS-MAP [17,35]. In [35], MDS was used to estimate locations in with anisotropic topology and complex terrain and to eliminate measurement error cumulation through iterative computations. In [17], the authors proposed a distributed and weighted MDS approach, dwMDS, thus avoiding the centralization shortcoming of the classical MDS-MAP approach.

(c) *Multihop localization approaches*: It is not always possible that a target trying to locate itself is in communication range with at least three anchor nodes because of the limitation of the transmission power. For that reason, mechanisms for multihop localization have been proposed to extend the localization process over a larger geographical extent.

One of the most popular methods is *Ad-Hoc Positioning System* (APS), proposed by Niculescu and Nath in [60]. APS aims at computing a range estimate between a target node faraway from anchors, assuming that a set of anchors is available in the network. The main idea of APS, similar to Ad hoc On-Demand Distance Vector (AODV) mechanism; it consists in flooding the network such that each anchor independently broadcasts a packet, called *beacon*, embedding its location and a hop-counter field initially set to one and increased in each new hop. Then, each target node identifies the shortest-path to each anchor node and tries to estimate its distance to it. Three *distance propagation methods* have been proposed: DV-HOP, Distance-Hop and Euclidian. DV-Hop is the only range-free method among all other propagation methods, which we describe in what follows.

The idea of DV-Hop is to compute the number of hops between any two anchors (A_i, A_j) and estimate the average 1-hop distance by dividing the sum of physical distances by the sum logical distances. More explicitly, each anchor estimates the average 1-hop distance by dividing the sum of its distance to other anchors by the sum of hop counts to those anchors. The DV-Hop steps can be summarized as follows:

– *Node update phase*: When a target N_i receives a beacon from an anchor, it maintains the record (X_i, Y_i, h_i) for each anchor A_i, where (X_i, Y_i) represents the location of the anchor, and h_i, the number of hops from N_i to that anchor A_i. At the end of this

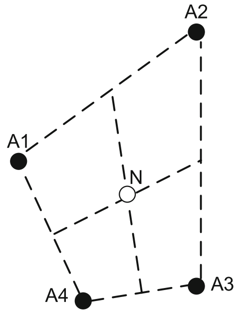


Fig. 9 Illustration of centroid localization technique

step, the target knows about the locations of anchor nodes and the hop counts to reach them.

– *1-hop distance estimation phase*: When an anchor node receives the locations and hop counts to other anchors, it calculates the estimated average 1-hop distance, referred to as *correction factor* c_i, expressed as follows:

$$c_i = \frac{\sum\left(\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}\right)}{\sum(h_i)} \tag{13}$$

Then, the anchor floods the network with the estimated 1-hop distance. A node that receives a correction, forwards it and then stops forwarding subsequent corrections.

– *Target node localization*: each target uses the correction sent from the closest anchor as the estimated 1-hop distance. It then multiplies the 1-hop distance by the hop counts to other anchors to estimate its physical distances to them. After getting distance estimates to at least three anchors, a target can use trilateration to approximate its location.

DV-Hop has the advantage of being simple and computationally-efficient, and also it does not depend on measurements error, for instance like when using RSS for distance estimation. However, DV-Hop is limited for use in isotropic networks to exhibit efficient behavior. In fact, for anisotropic environment, the average 1-hop distance will not be accurate as connectivity will be less correlated with range.

Gradient localization algorithm proposed by Nagpal et al. in [59] is another technique similar to DV-HOP in the sense that anchor nodes broadcast a message containing its location and hop count set to one, which increases from one hop to another. The difference with DV-Hop relies in the way how the 1-hop distance is calculated. The Gradient approach uses the formula given by Kleinrock-Silvester in [40] to estimate the 1-hop distance as the Euclidian distance covered by one radio hop expressed by:

$$d_{hop} = r\left(1 + \exp\left(-n_{local}\right) - \int_{-1}^{1} \exp\left(\frac{-n_{local}}{\pi}\left(\arccos t - \sqrt{1-t^2}\right)\right) dt\right) \tag{14}$$

where n_{local} represents the expected local neighborhood. The above estimation of 1-hop distance may result in additional localization errors.

Most of the above aforementioned localization methods assume an isotropic network topology with a statistically identical connectivity information for all directions. However, this assumption often does not hold in reality because of obstacles presence, radio irregularity

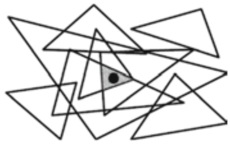


Fig. 10 Illustration of APIT localization technique

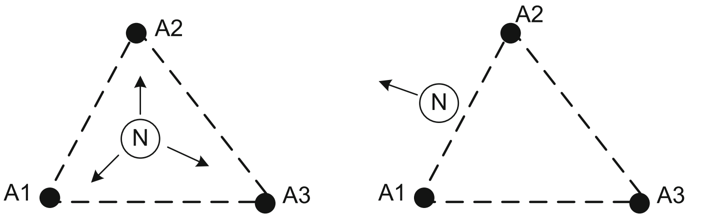


Fig. 11 Possible cases in point-in-triangle test. N denotes the target node, and A_i denotes the ith anchor. The *left figure* shows that if the target N is inside the triangle moves in any direction, it will get close to some anchors and far from others. In the *right figure*, if the target node N is outside the triangle moves in the indicated direction, it will get far to all anchors at the same time

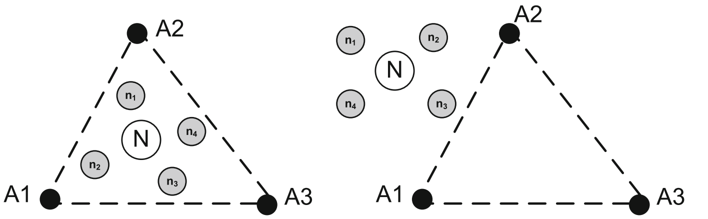


Fig. 12 Possible cases in approximate point-in-triangle test. N denotes the target node, and A_i denotes the ith anchor. The *left figure* shows that if the target N is inside the triangle, none of its neighbors is either close to or far from all anchors. In the *right figure*, if the target node N is outside the triangle, its neighbor n_1 indicates that is it further to all anchors than it

and the non-uniform node density. As a matter of fact, holes and complex shapes will be frequently present in the network topology. Several works have tried to deal with the anisotropic radio propagation issue including [44–46]. For instance, in [44] the authors have addressed the hole and complex shape problem by efficiently sub-devising the network, strategically placing anchor nodes, and finally selecting anchors placed on network boundaries with sufficient density. The method presented in [44] relies on (i) the use of geometrical concepts in order to construct the Voronoi diagram of anchors in the shape of triangles [85] and (ii) on rigidity theory in order to find the network layout and localize target nodes. A comparision between range-free anchor based techniques is shown in Table 6.

3.2.2 Anchor-Free Methods

Range-free anchor-free methods also called *Event-driven methods*, are schemes that exploit temporal and spatial properties of an event. For instance, an event may be the reception of an acoustic or a RF-signal. In [76], the authors argue that these methods can lead to a much higher accuracy as compared to the anchor-based ones. However, this benefit comes along with an implicit assumption that events can be precisely generated and propagated to a specific location at a specific time [85]. The well known anchor free methods are Walking-GPS, Lighthouse and Spotlight. Detailed descriptions are given in what follows.

Table 6 Comparison between range-free anchor based techniques

Localization methods	Accuracy	Overhead of computation	Overhead of communication	Number of anchors	Degree of scalability
Weighted centroid	≤Radio transmission range (R)	Depends on weights	Low, one hop	At least three	High
APIT	≤Radio transmission range (R)	Low	Low, one hop	Large	High
DV-hop	Low in anisotropic environments	Low	High	Low	Low
MDS-MAP	O(R)	High	High	Can be zero	Low

Walking-GPS: As its name reflects, walking GPS is a localization method consisting of a walking device/person carrying a Global Positioning System (GPS) which broadcasts periodically its/his position. The broadcasted data are used by unknown nodes to deduce their own locations. The hardware architecture of Walking-GPS encloses three main components:

– *A GPS device:* computes periodically its current position.
– *A mote connected to the GPS (GPS Mote):* broadcasts the coordinates got from the GPS device. For the sake of communication overhead reduction, Walking GPS system uses a local Cartesian coordinates system instead of the original GPS coordinates as these latter require 11 bytes for their representation.
– *Unknown nodes (Sensor Motes):* run a triangulation-based localization algorithm. This algorithm relies on the use of location information generated from two distinct sources: the GPS Mote and the settled nodes. Location information sent by settled nodes are used by unknown nodes that were unable to hear the data broadcasted by the GPS Mote.

The main advantage of the walking-GPS system is enabling by a single GPS device, the localization of an entire network with an acceptable accuracy (average localization errors are within 1 to 2 meters [75]).

Lighthouse: It is another anchor-free method proposed in [70]. As depicted in Fig. 13, this system consists of a constant rotating speed anchor node generating light signals. Its design assumes two idealistic hypothesis:

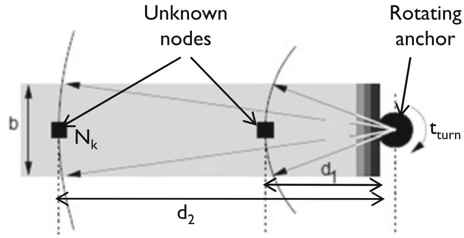
– A free-space optical channel between the light source and the unknown nodes.
– A parallel light beam.

While the beacon node is rotating, an unknown node N_k perceives the light signal during a period of time t_{beam}. t_{beam} depends on three parameters: the N_k's distance d from the rotation axis, the time t_{turn} required for a complete rotation and the light beam width b. t_{beam} is given by Eq. 15.

$$t_{beam} = \frac{\arcsin\left(\frac{b}{2d}\right) * t_{turn}}{\pi} \tag{15}$$

The location of the unknown node belongs to the cylinder with radius d centered at the lighthouse rotation axis. The main drawback of Lighthouse localization system is its idealistic assumptions. In fact, it is very difficult to ensure a parallel light beam. Ignoring such fact leads to inaccuracies. To cope with this drawback, the authors of [70] adapt their system by using two semiconductor laser modules and two rotating mirrors. Mirrors are mounted to ensure that the beam can be seen from any directions. Each laser module has a beam of width

Fig. 13 Lighthouse localization system [70]



b_i and angle orientations β_i,γ_i and δ_i where i = 1, 2. The resulting beam seen by the unknown node is approximately equal to:

$$b \approx b_1 + b_2 + \sqrt{d^2 + h^2} * (sin\beta_1 + sin\beta_2) \\ + h\,(tan\gamma_1 + tan\gamma_2) + d\,(sin\delta_1 + sin\delta_2) \tag{16}$$

where d is the unknown node distance to the lighthouse rotation axis while h is its height over the lighthouse center.

Spotlight: It is a localization system which employs an asymmetric architecture where all expensive and energy greedy operations are shifted from unknown nodes to a single powerful device. This device, called Spotlight is responsible of the computation of all unknown nodes locations. To do, Spotlight device flies over the network A and broadcasts light events $e(t)$ according to a predetermined event distribution function $E(t)$. When an unknown node detects an event, it reports back to the Spotlight device, its detection timestamps. This information will be mapped to a fixed position which will be also reported to the unknown node. As depicted in Fig. 14, Spotlight localization system supports three main functions dispatched between the Spotlight device and unknown nodes:

- A binary event detection function $D(e)$ set to true if an event e is detected and to false otherwise.
- An event distribution function $E(t)$ defining where in the network A an event was detected at time t.

$$E(t) = \{p | p \in A \wedge D(e(t, p)) = true\} \tag{17}$$

  where $e(t, p)$ is an event occurred at time t at the position $p \in A$.
- Localization Function $L(T_i)$ defining the localization algorithm having as inputs the timestamps Ti reported by the node i:

Spotlight localization system could be implemented in three possible instances:

1. *Point Scan:* In this instance, the Spotlight device is assumed generating light spots and moving with a constant speed s along a rectilinear line where unknown nodes lie.
2. *Line Scan:* In this instance, the Spotlight device is assumed generating a line of light events (e.g laser device) over a square shaped area A where unknown nodes are deployed. The Spotlight device is supposed scanning the area A by moving straightly along the x axis and then along the y axis.
3. *Area Cover:* Differently to Point Scan and Line Scan, the Area Cover supposes that Spotlight device is able to generate events covering a whole area at time (e.g video projector).

As it may be intuitively concluded, the Spotlight localization system requires that unknown nodes and the Spotlight device to be synchronized which may result in an increase in system cost.

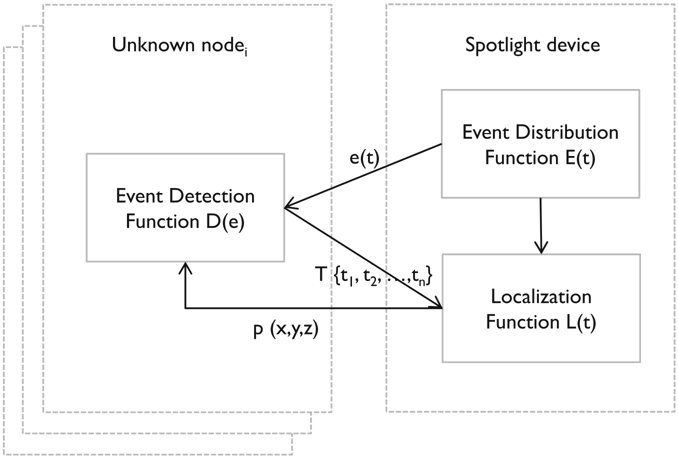
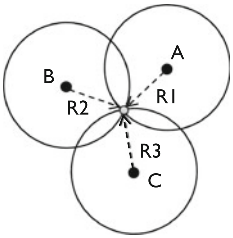


Fig. 14 Spotlight system architecture [76]

Fig. 15 Trilateration with noise-free measurements



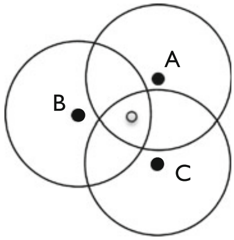
3.3 Geolocation Techniques

Geolocation is the act of finding unknown nodes locations based on some information pertaining to distances and angles measurements. Basically, there are two major classes of gelocation techniques: (1) *geometric geolocation techniques* (Table 7) based on geometric transformation of distance or angle information to estimate the location of the target; *lateration* and *triangulation* are the fundamental methods of this class, and (2) *refinement geolocation techniques* consist of mathematical methods aiming to reduce measurements' noise and to improve the accuracy of the estimated location. The following sections provide more details about both classes.

3.3.1 Geometric Geolocation Techniques

Lateration/Trilateration Lateration is the technique that uses distance information from anchor nodes to locate a target. In a 2D space, lateration involves the determination of the location of an unknown node as the intersection point of three circles centered in three non-collinear anchors (A, B and C), given that distances R_1, R_2 and R_3 (i.e. circle radii) between the node and anchors are known. This technique is referred to as *trilateration*. In a 3D space, there is a need of at least four anchor nodes to determine the location of a target node. In the ideal case, trilateration assumes that distance measurements are precise and noise-free, as depicted in Fig. 15. Such a situation is not usually true as errors and inaccuracies most likely occur. These inaccuracies prevent circles intersection to be an exact point Fig. 16 and makes

Fig. 16 Trilateration with noisy measurements



the localization process more challenging. In this case, one possible solution to estimate the target location is to use Maximum Likelihood algorithm [25] or the least squares optimization method. More details about least-squares methods will be presented in Sect. 3.3.2.

The trilateration process for inferring the target location is achieved by solving the linear equation system in Eq. (18),

$$(x_i - x)^2 + (y_i - y)^2 = {d_i}^2 \tag{18}$$

where x and y are the coordinates of the unknown node, and (x_i, y_i) are the coordinates of at least three anchors N_i, $i \in 1, 2, 3$ involved in the localization. The linear equation system would give an exact and unique solution if the circles intersect in one point, in an ideal noise-free environment. As shown in [24], this system can be linearized by subtracting Eq. (18) for an anchor N_i from the equivalent expression of anchor N_1, it results:

$${d_1}^2 - {d_i}^2 = {x_1}^2 + {y_1}^2 - {x_i}^2 - {y_i}^2 + 2x(x_i - x_1) + 2y(y_i - y_1) \tag{19}$$

The system can be written in matrix notation:

$$HX = B \tag{20}$$

where $X = [x, y]^T$ represents the target node location,

$$H = \begin{bmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \\ \vdots & \vdots \\ x_n - x_1 & y_n - y_1 \end{bmatrix} \tag{21}$$

and

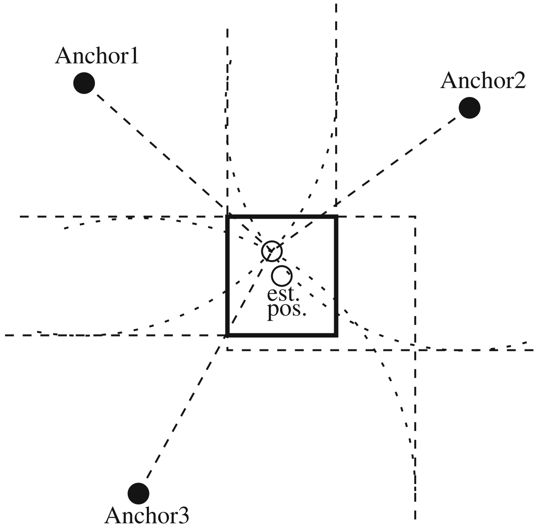
$$B = \frac{1}{2} \begin{bmatrix} \left(d_1^2 - d_2^2\right) + \left(x_2^2 + y_2^2\right) - \left(x_1^2 + y_1^2\right) \\ \left(d_1^2 - d_3^2\right) + \left(x_3^2 + y_3^2\right) - \left(x_1^2 + y_1^2\right) \\ \vdots \\ \left(d_1^2 - d_n^2\right) + \left(x_n^2 + y_n^2\right) - \left(x_1^2 + y_1^2\right) \end{bmatrix} \tag{22}$$

Thereby, the target location can be estimated as solution of [Eq. (20)]:

$$X = (H^T H)^{-1} H^T B \tag{23}$$

Bounding-Box (Min–Max) The bounding-box (also known as min–max) algorithm is another computationally-efficient alternative to trilateration that relies on the intersection of rectangles instead of circles to estimate the location of an unknown node. The main idea is to draw a bounding box for each anchor node using its location and distance estimate, then to determine the intersection of these rectangles. The location of the target node is estimated

Fig. 17 Bounding-box (Min–Max) [42]



as the center of the intersection rectangle. Figure 17 illustrates the bounding-box method for a target node based on the distance estimate of three anchor nodes. The min–max method provides a solution very close to the ideal solution obtained through trilateration, with much less computation requirements. Formally, the bounding box pertaining to an anchor N_i is constructed by subtracting its distance estimate d_i from its location $[x_i, y_i]$

$$[x_i - d_i, y_i - d_i] \times [x_i + d_i, y_i + d_i] \tag{24}$$

The intersection of the bounding boxes is computed by taking the maximum of all coordinate minimums and the minimum of all maximums:

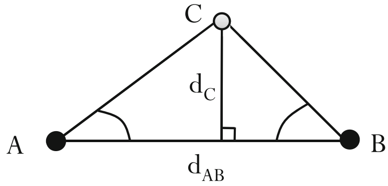
$$[max(x_i - d_i), max(y_i - d_i)] \times [min(x_i + d_i), min(y_i + d_i)] \tag{25}$$

It can be noticed that the computation cost with bounding-box method is much smaller than that with trilateration, at the expenses of a bit lower accuracy.

Triangulation Triangulation is a geolocation method that exploits triangles properties in order to determine unknown nodes locations. Being different from trilateration, triangulation is based on angle measurements to estimate the location of an unknown node rather than measuring distances to the unknown node (i.e. trilateration). In particular, it is typically based on AOA measurements from two anchor nodes in a 2D space. In a 3D space, triangulation would be possible if another measurement of azimuth is available. Two cases may arise with triangulation: (1) the distance between the two anchor nodes is known, (2) the distance between the two anchor nodes is unknown.

In the first case, if we consider two anchor nodes A and B separated by a distance d_{AB}, C is the target node, and $\widehat{CAB}$ and $\widehat{ABC}$ are two known angles, then the location of the unknown node C completing the constitution of the triangle ABC is derived using the trigonometry laws of sines and cosines (also known as AL-KASHI theorem). Node C will be located at

Fig. 18 Triangulation (first case)



distance d_C at the perpendicular to line (A, B) as shown in Fig. 18 such that:

$$d_C = \frac{d_{AB} \sin\left(\widehat{CAB}\right) \sin\left(\widehat{ABC}\right)}{\sin\left(\widehat{CAB} + \widehat{ABC}\right)} \tag{26}$$

In the second case, assuming that each anchor N_i is able to measure the AOA θ_i of the signal transmitted from the target node, then it is possible to write:

$$(x_i - x)\sin(\theta_i) = (y_i - x)\cos(\theta_i) \tag{27}$$

and, in matrix form:

$$HX = B \tag{28}$$

where

$$H = \begin{bmatrix} -sin\,(\theta_1) & cos\,(\theta_1) \\ -sin\,(\theta_2) & cos\,(\theta_2) \\ \vdots & \vdots \\ -sin\,(\theta_n) & cos\,(\theta_n) \end{bmatrix} \tag{29}$$

and

$$B = \begin{bmatrix} y_1 cos\,(\theta_1) - x_1 sin\,(\theta_1) \\ y_2 cos\,(\theta_2) - x_2 sin\,(\theta_2) \\ \vdots \\ y_n cos\,(\theta_n) - x_n sin\,(\theta_n) \end{bmatrix} \tag{30}$$

The solution is given by:

$$X = \left(H^T H\right)^{-1} H^T B \tag{31}$$

Multilateration or (Hyperbolic localization) Multilateration, also known as *hyperbolic localization* [34], is a geolocation technique based on the estimation of distances using TDOA measurements. Multilateration typically refers to locating an emitter node with unknown location by measuring the TDOA of the signal it emits to three of more anchor nodes. It may also refer to locating a receiver node that measures the TDOA of a signal transmitted from three or more synchronized anchor nodes. Multilateration is different from trilateration, which relies on absolute (real) distance measurements between the target node and anchor nodes, whereas multilateration exploits the differential distance between anchor nodes. We also note that multilateration differs from triangulation, which is based on at least two-angle measurements. The nomination of multiletration as hyperbolic localization is because the possible location of a target node lies in a hyperbola as a result of considering the distance difference (i.e. TDOA) instead of absolute distances. In fact, the hyperbola is defined as the curve such that the difference of the distance from any point on the hyperbola to the two foci is a constant $2a$, which is the distance between its two vertices [52]. If the distance between the two foci is $2c$, then the hyperbola equation is expressed as:

$$\sqrt{(x+c)^2+y^2}-\sqrt{(x-c)^2+y^2}=2a \tag{32}$$

It can also be written in the following equivalent form:

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \tag{33}$$

where $b=\sqrt{c^2-a^2}$.

For a TDOA measurement t_i that pertains to a difference of distances D_{ij} between the target node and the anchors N_i and N_j, the hyperbola equation is:

$$d_i-d_j=D_{ij} \tag{34}$$

which is equivalent to Eq. (32). d_i and d_j represent the distance between the target node and anchors N_i and N_j, respectively. The relation between the distances and the parameters of the hyperbola is given in [52]. Similarly to trilateration, due to measurements inaccuracy, the intersection of hyperbolas will not result into a single point. The same optimization approaches as those used in case of trilateration can also be used in this case. In particular, Eqs. (20) and (23) can still be applied for the following matrix notation of the problem:

$$X=[x,y,d_1] \tag{35}$$

where

$$H=\begin{bmatrix} x_2-x_1 & y_2-y_1 & d_2 \\ x_3-x_1 & y_3-y_1 & d_3 \\ \vdots & \vdots & \vdots \\ x_n-x_1 & y_n-y_1 & d_n \end{bmatrix} \tag{36}$$

and

$$B=\frac{1}{2}\begin{bmatrix} \left(x_2^2+y_2^2\right)-\left(x_1^2+y_1^2\right)-d_2^2 \\ \left(x_3^2+y_3^2\right)-\left(x_1^2+y_1^2\right)-d_3^2 \\ \vdots \\ \left(x_n^2+y_n^2\right)-\left(x_1^2+y_1^2\right)-d_n^2 \end{bmatrix} \tag{37}$$

We note that d_1 refers to the distance between the target node and the first anchor N_1 and that this distance is also unknown, thus is included in the unknown vector X. Using the least-squares optimization method as in Eq. (23), a solution can be determined for this system in noisy environments. Other techniques such as Extended Kalman Filter can also be used.

The Closest-Neighbor Algorithm It is also known as proximity-based localization. In the Closest-Neighbor algorithm (CN), the location of the unknown node is simply confused with the location of the closest anchor. The algorithm proceeds as follows: given a group of anchors, in order to locate a particular node s_n, a distance measurement is performed by each anchor N_i. Let d_i be the set of all measured distances. The location of s_n will be determined as the location of the anchor having the minimum d_i.

Table 7 Comparison between Geometric geolocation techniques

Geolocation methods	Accuracy	Complexity of computation	Requirements
Bounding-box	Coarse-grained	$O(n)$ comparisons	At least two anchors
Triangulation	Fine-grained if non-noisy measurements	$O(n)$ multiplications	At least two anchors, angle measures
Trilateration	Fine-grained if non-noisy measurements	$O(n)$ multiplications	at least three anchors
Multilateration	Fine-grained if non-noisy measurements	$O(n)$ multiplications	At least three anchors, synchronization between anchors
The closest-neighbor	Coarse-grained	Zero	At least one anchor

3.3.2 Refinement Geolocation Techniques

Least-Squares method The main objective of Least Square (LS) method in localization consists in minimizing the effect of distance errors on the estimated location. In fact, when using trilateration, for example, the intersection of circle does not result in a single point since the estimated distances are different from real distances. The LS consider the localization problem as an optimization problem and finds the optimal location that minimizes the square errors.

Least Squares (LS) models the range errors as random variables η_i affecting the measured distances ($D = [d_1, \ldots, d_n]^T$). Range errors take into account the measurements errors resulted from channel noise, and NLOS errors resulted from blocks presence in the direct path [63]. Thus, Eq. (18) can be written as:

$$d_i = f_i(X) + \eta_i \tag{38}$$

where $X = (x, y)$ is the coordinates vector of the unknown node, $f_i(X)$ is the distance computation function $\sqrt{(x - x_i)^2 + (y - y_i)^2}$ and (x_i, y_i) are the coordinates of the anchor nodes N_i, $i \in 1, 2, \ldots, N$ involved in the localization.

The LS problem searches to determine the optimal unknown node coordinates $\hat{X} = (\hat{x}, \hat{y})$ by finding the argument minimizing the cost function $H(X)$:

$$\hat{X} = arg\, \min_{x} \{H(X)\} \tag{39}$$

where

$$H(X) = \sum_{i=1}^{N} [d_i - f_i(X)]^2 = [D - F(X)]^T [D - F(X)] \tag{40}$$

and $F(X) = [f_1(X), \ldots, f_n(X)]^T$. The solution to Eq. (39) can be obtained by setting the gradient of $H(X)$ to zero, where the partial derivative of $H(X)$ is:

$$\nabla H(X) = -2\,[D - F(X)]^T\, \nabla F(X) \tag{41}$$

LS method presents one major limitation. In fact, when measuring the optimal location, LS takes equitably all distance measurements computed by all the N anchors. However, these measurements have different accuracy degrees. Thus, using all measured data does not necessarily lead to the derivation of the best possible unknown node location. A better approach consists in applying a weight to each distance measurement in order to mitigate largely erroneous data. This constitutes the idea behind the Weighted Least Squares (WLS) method proposal. WLS assumes that the distance measurement d_i is corrupted by a white Gaussian noise having a standard deviation equals to σ_i, i.e. $\eta_i \sim$ Norm(0,σ_i) where η = $[\eta_1, \ldots, \eta_n]^T$ and $E[\eta\ \eta^T]$ = R, the best fit is when the following cost function $H(X)$ is minimized [24]:

$$\begin{aligned} H(X) &= \sum_{i=1}^{N} R_{i,i}^{-1}\, [d_i - f_i(X)]^2 \\ &= [D - F(X)]^T\, R^{-1}\, [D - F(X)] \end{aligned} \tag{42}$$

Similarly to LS, the optimal location can be calculated by setting the gradient of $H(X)$ to zero:

$$\nabla H(X) = -2\, [D - F(X)]^T\, R^{-1} \nabla F(X) = 0 \tag{43}$$

In Eq. (43), R is a diagonal matrix defining the different weights.

RWGH Algorithm The Residual Weighting alGoritHm (RWGH [63]) has been proposed to overcome the limitation of the Least-Square method discussed previously in Sect. 3.3.2. In fact, the estimation of unknown node location in RWGH is not computed based on all collected distance measurements. Rather, RWGH derives it by estimating preliminary intermediate locations computed over multiple sub-sets of the measured data. The final RWGH output consists in a weighted sum of all estimated locations. RWGH algorithm proceeds according to the following steps. Given a set $\{D_i\}_{3 \le i \le M}$ of distance measurements representing each the distance separating the unknown node and an anchor node i, RWGH computes the set S of all possible distance measurements combinations. where $S = \{\binom{M}{i}\}_{3 \le i \le M}$ and $\binom{M}{i}$ denotes the set of all possible combinations of i measurements selected from a total of M measurements.

Each combination in RWGH is referred by an index $S_k | k = 1, 2, \ldots, N_c$ where $N_c = card(S) = \sum_{i=3}^{M} \frac{M!}{i!(M-i)!}$. Subsequently, the Least Square method is applied to each combination S_k in order to determine the best unknown node location coordinates $\hat{X}_k = (\hat{x_k}, \hat{y_k})$ which minimizes the residual. $\hat{X}_k$ is considered as an intermediate estimate of the unknown node coordinates and it is given by:

$$\hat{X}_k = arg\ \underset{X}{min} \left\{ \sum_{i \in S_k} \left(d_i - \sqrt{(x - x_i)^2 + (y - y_i)^2} \right)^2 \right\} \tag{44}$$

The minimal residual corresponding to the combination S_k is then normalized by the number of elements of this latter. This quantity is referred by $\widetilde{R_{es}}(\hat{X}_k, S_k)$. The final estimate of the coordinate vector $\hat{X} = (\hat{x}, \hat{y})$ is expressed as the weighted sum of the intermediate estimates, as shown in Eq. 45.

$$\hat{X} = \frac{\sum_{k=1}^{N_c} \hat{X}_k * \left(\widetilde{R_{es}} \left(\hat{X}_k, S_k \right) \right)^{-1}}{\sum_{k=1}^{N_c} \left(\widetilde{R_{es}} \left(\hat{X}_k, S_k \right) \right)^{-1}} \tag{45}$$

In the RWGH algorithm, it is clear that if M increases N_c goes larger. As for each sub-set $S_{k_{1\leq k\leq N_c}}$, a least-square method needs to be applied, the computational complexity of RWGH becomes quite high [36].

4 Localization Accuracy Metrics

Accuracy is the fundamental criterion that reflects the goodness and the performance of a localization technique. It is defined by how close the estimated and the actual locations are. In other words, accuracy reflects the amount of errors in the estimated locations. Several factors impact the resolution of accuracy such as distorted measurements, memory and computation constraints, dynamic environment changes, path loss effects and most importantly errors accumulation and propagation. Nevertheless, despite its importance, accuracy is not the overriding goal of a good localization technique as this is application-dependent [57]. For instance, in fire-alerting system, fires may be localized with an accuracy of some meters however automated guided vehicles requires much more accuracy in the range of few centimeters. In the literature, several accuracy metrics have been proposed. They can be classified into three categories: distance-based, position-based and area-based as depicted in Fig 19. The following sub-sections provide a bird's eye view on the most relevant metrics of these categories.

4.1 Distance-Based

Distance-based accuracy metrics measure the localization accuracy based on the information of estimated distances.

4.1.1 Average Relative Deviation (ARD)

ARD [26] represents the average deviation ratio between estimated and actual distances of two nodes i and j. ARD metric takes into account both short-range and long-range errors. In fact, the deviation ratio is averaged over all possible nodes pairs, and not only adjacent ones. It is expressed as:

$$ARD = \frac{2}{n\,(n-1)} \sum_{i<j}^{n} \frac{\left|\hat{d_{ij}} - d_{ij}\right|}{min\left(\hat{d_{ij}}, d_{ij}\right)} \tag{46}$$

4.1.2 Global Energy Ratio (GER)

This metric has been proposed in [65] in order to measure how well the estimated network layout derived from estimated inter-nodes distances matches the actual network layout. In other words, GER quantifies the estimated layout error and it is calculated using the following equation:

$$GER = \frac{1}{n\,(n-1)\,/2} \sqrt{\sum_{i=1}^{n} \sum_{j=i+1}^{n} \left(\frac{\hat{d}_{ij} - d_{ij}}{d_{ij}}\right)^2} \tag{47}$$

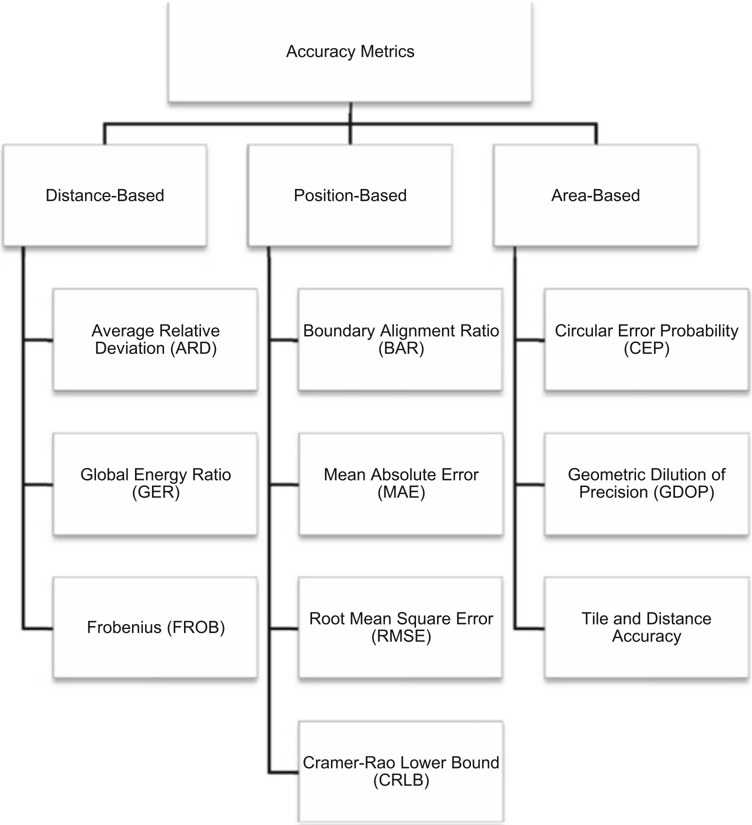


Fig. 19 Accuracy metrics

where n denotes the number of nodes in the network, $\hat{d}_{ij}$ and d_{ij} are respectively the estimated and the actual distances between the two nodes i and j. The authors of [20] argue that GER metric is appropriate to compare the qualities of layouts obtained by different localization algorithms for graphs of the same size, nevertheless it is not well-suited to compare between those of different graph sizes.

4.1.3 Frobenius (FROB)

Similarly to GER metric, FROB metric [20] has been proposed to verify if the estimated network layout created by the localization algorithm matches the actual one. Consider a network of n nodes where $\hat{d}_{ij}$ and d_{ij} are respectively the estimated and the actual distances between the two nodes i and j, FROB is equivalent to the Frobenius norm of the matrix M whose entries are:

$$M_{ij} = \frac{\hat{d}_{ij} - d_{ij}}{n} \tag{48}$$

In other words, FROB metric computes the normalized error of the global estimated network layout.

$$FROB = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(M_{ij}\right)^{2}} = \sqrt{\frac{1}{n^{2}}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\hat{d}_{ij} - d_{ij}\right)^{2}} \tag{49}$$

4.2 Position-Based

Unlike distance-based accuracy metrics, metrics belonging to this category evaluate how accurate the estimated positions are.

4.2.1 Mean Absolute Error (MAE)

The MAE metric measures the average distance (known also as the residual or deviation) between estimated and actual position coordinates. Given a network of n nodes, where for each node of index $i_{1 \leq i \leq n}$, the actual position coordinates vector X_i is known, MAE is equal to:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left(\left\|X_i - \hat{X}_i\right\|\right) \tag{50}$$

where $\hat{X}_i$ is the estimated position coordinates vector for the given node i.

4.2.2 Root Mean Square Error (RMSE)

RMSE is commonly used in statistic and it is roughly similar the MAE metric as both of them determine the residual between estimated and actual data. RMSE is used in localization context in order to quantify the residual between estimated and actual position coordinates vectors and this is by computing the standard deviation between the two vectors.

$$RMSE = \sqrt{E\left[\left(X_i - \hat{X}_i\right)^{2}\right]} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \hat{X}_i\right)^{2}} \tag{51}$$

In both RMSE and MAE, values near to zero reflect high accuracy resolution.

4.2.3 Cramer-Rao Lower Bound (CRLB)

The CRLB metric defines the lower bound for the variance of any unbiased estimator. In theory, this variance is at least as high as the inverse of the Fisher Information Matrix (FIM). If the variance is equal to the CRLB, then the estimation is considered to be optimal. CRLB is computed as follows:

$$Var\left(\hat{\theta}\right) \geq \left(-E\left[\frac{\partial^{2} log\ p\left(r, \theta\right)}{\partial \theta^{2}}\right]\right)^{-1} \tag{52}$$

where, θ is a column vector of parameters to be estimated, $p\left(r, \theta\right)$ is the probability density function of a random variable r, and $\hat{\theta}$ is an estimator of θ. In typical localization problem, θ represents the vector of the actual position coordinates of sensor nodes, whose locations are to be estimated.

4.2.4 Boundary Alignment Ratio (BAR)

This metric has been also proposed by [20] in addition to the FROB metric. While the latter metric compares the accuracy of the estimated network layout globally, the BAR metric estimates how well the estimated positions of nodes that sit on the boundary match their actual positions. The Bar metric is evaluated as follows:

$$BAR = \frac{1}{|S|} \sum_{\hat{x} \in \hat{S}} \left(\hat{x} - x\right)^2 \tag{53}$$

where S and $\hat{S}$ are respectively the actual and the estimated set of nodes that sit on the boundary of the network, $|S|$ is the size of S and x is the closest node in S to $\hat{x}$.

4.3 Area-Based

Area-based techniques represent another way to quantify the localization accuracy and they are characterized by the size of the area where the unknown node is likely to be. The smaller the area is, the better is the accuracy. In what follows, we define the most relevant techniques.

4.3.1 Circular Error Probability (CEP)

This metric defines the radius d of a circular area where 50 % of the residual $\|X_i - \hat{X}_i\|$ are within that area. More formally, the circular error probability (CEP) is expressed as:

$$CEP = d \setminus P\left(\frac{\left|\left\|X_i - \hat{X}_i\right\| \leq d\right|}{|S|}\right) = 0.5 \tag{54}$$

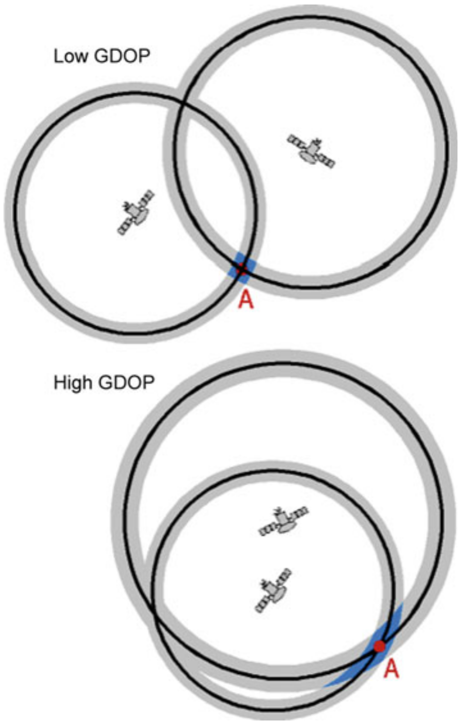
where $\|X_i - \hat{X}_i\|$ denotes the residual between an estimated and actual position coordinates vector for a given node i and S is the set containing all $\|X_i - \hat{X}_i\|$ values (i.e $S = \{\|X_i - \hat{X}_i\|\}_{1 \leq i \leq n}$). Based on the previous formula, we notice that when the radius d increases, the accuracy decreases as the area becomes larger.

4.3.2 Geometric Dilution of Precision (GDOP)

GDOP metric is related to GPS localization context. Recall that GPS system infers unknown node location using 3-D lateration which consists in finding where spheres intersect. If measured spheres' radius are corrupted by noise then spheres will intersect in an oddly-shaped area and not in an exact point. GDOP metric measures how large this area is. As depicted in Fig. 20, the size of this area is impacted by the geometric scattering of satellites chosen in the 3-D lateration process. If the satellites are close to each other, the geometry is bad and the GDOP value is high however if they are far apart, the geometry is good and the GDOP value is low. Such result is due to the fact that satellites that are close to each other provide less information than those that are widely separated. Thereby, higher the GDOP is, worse is the accuracy. GDOP metric is given by [86]:

$$GDOP = \sqrt{tr\left(G^T G\right)^{-1}} \tag{55}$$

Fig. 20 Satellites geometric scattering impact [74]



where G is equal to:

$$G = \begin{pmatrix} \frac{x_1-x}{R_1} & \frac{y_1-y}{R_1} & \frac{z_1-z}{R_1} & 1 \\ \frac{x_2-x}{R_2} & \frac{y_2-y}{R_2} & \frac{z_2-z}{R_2} & 1 \\ \frac{x_3-x}{R_3} & \frac{y_3-y}{R_3} & \frac{z_3-z}{R_3} & 1 \\ \frac{x_4-x}{R_4} & \frac{y_4-y}{R_4} & \frac{z_4-z}{R_4} & 1 \end{pmatrix}$$

where x, y and z denote the position of the GPS receiver, x_i,y_i and z_i denote the position of satellite i and R_i denotes the distance of the GPS receiver to the satellite i. GDOP metric can be used in lateration geolocation technique in order to find the best combination of reference nodes.

4.3.3 Tile and Distance Accuracy

This metric was proposed by [21] and it is used when the reported location is described as a set of small discrete tiles rather than a single point or a single area. Tile accuracy refers to the percentage of times the localization technique is able to return the true tile containing the unknown node. The drawback of this metric is that sometimes, the true tile is close to the returned set, nonetheless the localization accuracy is reported as bad. To overcome this shortcoming, [21] has proposed Distance accuracy metric where accuracy is quantified by the distance between the true tile and tiles in the returned area.

5 Discussions and Future Directions

5.1 Discussions and Lessons

Localization has been attracting a lot of attention and represents a key research area in the community of the CPS, as it relates events to their locations in the surrounding environment. This paper fills a gap by presenting an attempt to provide a global and unified taxonomy of fundamental concepts of localization in CPS.

In the first part of this paper, we synthesized the literature and proposed a global taxonomy of localization concepts. We proposed a classification that helps researchers to have a complete knowledge of the different localization approaches and paradigms in CPS, while deeply analyzing their advantages and limitations. This is indeed very useful taking into account the diversity and the number of localization concepts and their applications. In the second part of this paper, we presented a thorough review of fundamental localization techniques, and we have reviewed the most relevant research works for each category. Finally, the last part of this review paper was devoted to localization accuracy, which is a key criterion to differentiate and compare between localization techniques.

It appears that the choice of an adequate localization technique for a particular cyber-physical application is a complex task as it is inherently dependent on several factors including, but not limited to: (i) *the desired solution accuracy*, which may vary from a few centimes to several meters. Range-free techniques such as Centroid may be sufficient for accuracy-tolerant systems (ii) *solution cost*: in fact additional hardware, such as ultrasonic devices, would produce good accuracy but this comes with an additional cost that may be not suitable for large-scale applications, (iii) *solution scale*, localization in large-scale systems is more challenging than that with small-scale systems. Distributed approaches would be more suitable for large-scale systems than centralized approaches. (iv) *environment*: the requirements of indoor localization are different from those of outdoor location systems, as presented in Section 2. In summary, the localization system designer needs to make a good balance between the different requirements of the target solution.

Furthermore, we observe that among all presented localization techniques, RSS-based localization is a range-based technique that represents a key method in resource-constrained low-cost devices, such as sensor nodes, for two main reasons: (1) it relies on built-in wireless transceivers and thus does not require any additional hardware as in other techniques such as ultrasound devices with time-based approaches, directional antennas in angle based approaches, (2) induces a low computational complexity as compared to other techniques, as it does not require intensive signal processing computation as required by image-based or sound-based localization methods. The complexity of the fingerprinting phase can be overcome with the recent techniques that perform distance to RSS mapping on runtime [7,14,33,49,79]. Deployment complexity represents one challenge among several others in what concerns the real-world utilization of localization mechanisms. The next section summarizes the main practical issues pertaining to localization.

5.2 Real-World Challenges

Building a robust localization system in real-world is a very hard problem as it encompasses several practical challenges. In fact, as mentioned above, the localization system designer have to ensure a tradeoff between several antagonistic metrics including system cost, energy effectiveness, ease of calibration and deployment, and accuracy. The achievement of this tradeoff is not obvious and wraps-up serious challenges. We classify these challenges into four categories:

– *Cost-effectiveness*: To ensure a high localization accuracy level, the underlying localization mechanism must incorporate extra sophisticated hardware per network nodes. Nonetheless, this may induce a heavy burden on the system cost in addition to energy dissipation, since cyber-physical devices, such as sensor nodes, are generally battery powered and massively deployed.
– *Measurements errors*: RF and acoustic signals are the main vehicles of localization measurements. These signals are inherently unreliable as they may get distorted. For instance, most of RF-signals have irregular propagation patterns induced by the environment conditions (namely pressure and temperature) and the random multi-path effects including reflection, refraction, diffraction and scattering. These phenomena result from the obstruction of physical objects during the signal propagation. On the other side, acoustic signals are also subject of distortion caused namely by the environment conditions and echo presence. Errors in measurement make the localization process more challenging as it is needed to filter the measurement noise out to improve localization accuracy.
– *Deployment complexity*: Typical localization methods, such as time-based and RSS-based methods, usually require a pre-deployment configuration process. For instance, in TOA-based technique, transceivers of different nodes must be accurately synchronized before starting localizing objects. Also, RSS-based techniques need to be pre-calibrated before their use. This calibration requires complex, tedious, labor-intensive, time-consuming, human-based and offline environment profiling phase. Such cumbersome pre-deployment phase is a major handicap constraining the wide adoption of the localization system and its practical use. This operation becomes even more complex as environment changes will compromise the caliber. To cope with this shortcoming, calibration needs to be automated and made environment adaptive. It is also important to pre-implement self-configuration mechanisms in order to cope with network dynamics (e.g. due to nodes failures). Furthermore, one of the most challenging problems pertaining to deployment of a localization system is to determine the optimal number of anchor nodes to be deployed in addition to their placement pattern. According to [23], two main constraints must be addressed for strategic placement patterns. First, it is important to maximize the coverage area while minimizing the number of anchor nodes. Second, the system tuned with the selected pattern has to (1) offer an acceptable accuracy degree and to (2) avoid interference between adjacent nodes. By skimming the state of the art, no general guidelines are available.
– *Security*: Securing localization mechanisms is essential for certain critical applications to protect the localization system against malicious attacks that may compromise the application security [13,43,69]. In general, these attacks threaten the integrity, the confidentiality and the availability of the location information. For instance, RSS-based localization techniques are vulnerable to Signal Strength attacks. Indeed, the RSS of an anchor or a target node can be attenuated or amplified by placing an absorbing or a reflecting material around the node [12]. Furthermore, because of the openness of CPS, they might be exposed to spoofing attacks [13] where a malicious node pretends to be a legitimate node and thus can inject undesirable traffic in the system. Another common attack is wormhole attack where malicious node sniffers transmitted packets at one location, sends them to another malicious node placed in another location, which in its side replays them locally [69]. This would make the localization system to operate incor-

rectly. All these challenges must be addressed before a real-deployment, in particular for applications where security is a main concern.

5.3 Open Issues

Throughout this review paper, we have presented a representative sample of the vast array of research works in the localization area. In spite the high number of works in the literature, there are still several challenges to be addressed in the future to meet the ever evolving nature of future cyber-physical networks. In what follow, we enumerate, without being comprehensive, some challenging ideas and research trends in the localization arena:

1. *Localization Data Fusion*: The predominant localization techniques basically rely on a single type of information (e.g. radio or sound or image) to localize nodes. In future CPS, the need for heterogeneous data fusion techniques for location estimation is a promising research area. Hybrid localization techniques should use different measurement data sources and collaborate together to estimate the location of unknown nodes. This is rather useful in robotic applications, where different sensor types are available.
2. *Zero Profiling Localization*: Most of radio-based localization techniques rely on offline fingerprinting methods and profiling approaches to characterize the environment. Such paradigm is not suitable for dynamic systems where the characteristics of the environment changes over time. There is need to develop more sophisticated and practical approaches for RSS-based localization without need to perform offline profiling of the environments. There has been some proposed approaches [9,48,50], but there still room for major improvement of these techniques.
3. *Novel Radio-based Localization Mechanisms*: Radio-based localization is very appealing, in particular for low-cost solutions, as it does not require additional hardware. The major trend has been the use of the RSS information to infer distances to and then locations of unknown nodes. However, RSS is known to be highly variable and does not map well with distance. In this regards, other link quality metrics such as LQI, SNR and F-LQE [3] could be considered to propose new radio-based localization mechanisms that provide better accuracy than RSS-based techniques. Another trend would be to use other link quality metrics to improve RSS-based localization.
4. *Benchmarking Localization Methodology* : One challenge that needs further research is the devising of a benchmarking methodology enabling objective experimental validation of and fair comparison between state-of-the art localization solutions. If such methodology is provided, the prototyping of new localization solutions would become easier.

Acknowledgments This work is funded by the R-Track project [72] under the grant 8-INF-2008 of the National Plan for Sciences and Technology (NPST), managed by the Science and Technology Unit of Al-Imam Mohamed bin Saud University and by King AbdulAziz Center for Science and Technology (KACST).

References

1. Ali-Loytty, S., Perala, T., Honkavirta, V., & Piche, R. (2009). Fingerprint kalman filter in indoor positioning applications. In *Control applications, (CCA) intelligent control (ISIC)* (pp. 1678–1683). doi:10.1109/CCA.2009.5281069.

2. Amundson, I., & Koutsoukos, X. D. (2009). A survey on localization for mobile wireless sensor networks. In *Proceedings of the 2nd international conference on mobile entity localization and tracking in GPS-less environments, MELT'09* (pp. 235–254). Berlin: Springer. http://dl.acm.org/citation.cfm?id=1813141.1813162.
3. Baccour, N., Koubâa, A., Youssef, H., Jamâa, M. B., do Rosário, D., Alves, M., & Becker, L. B. (2010). F-lqe: A fuzzy link quality estimator for wireless sensor networks. In *EWSN* (pp. 240–255).
4. Bahl, P., & Padmanabhan, V. N. (2000). Radar: An in-building rf-based user location and tracking system. In *INFOCOM* (pp. 775–784).
5. Barsocchi, P., Lenzi, S., Chessa, S., & Giunta, G. (2009). Virtual calibration for rssi-based indoor localization with ieee 802.15.4. In *IEEE international conference on communications (2009). ICC '09* (pp. 1–5). doi:10.1109/ICC.2009.5199566.
6. Ben Jamâa, M., Kouba, A., & Kayani, Y. (2012). Easyloc: RSS-Based localization made easy. *Procedia Comput Sci., 10*, 1127–1133. http://dx.doi.org/10.1016/j.procs.2012.06.160.
7. Bernardos, A. M., Casar, J. R., & Tarrıo, P. (2010). Real time calibration for rss indoor positioning systems. East (September), 15–17.
8. Blumenthal, J., Grossmann, R., Golatowski, F., & Timmermann, D. (2007). Weighted centroid localization in zigbee-based sensor networks. In *IEEE international symposium on intelligent signal processing (WISP 2007)* (pp. 1–6). doi:10.1109/WISP.2007.4447528.
9. Bolliger, P. (2008). Redpin—adaptive, zero-configuration indoor localization through user collaboration. In *Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments, MELT '08* (pp. 55–60). New York, NY: ACM. doi:10.1145/1410012.1410025.
10. Boukerche, A., Oliveira, H. A. B. F., Nakamura, E. F., & Loureiro, A. A. F. (2007). Localization systems for wireless sensor networks. *IEEE Magazine of Wireless Communications*, *14*(6), 6–12. doi:10.1109/MWC.2007.4407221.
11. Bulusu, N., Heidemann, J., & Estrin, D. (2000). GPS-less low cost outdoor localization for very small devices. *IEEE Personal Communications Magazine, 7*(5), 28–34.
12. Chen, Y., Kleisouris, K., Li, X., Trappe, W., & Martin, R. P. (2006). The robustness of localization algorithms to signal strength attacks: A comparative study. In *DCOSS'06* (pp. 546–563).
13. Chen, Y., Trappe, W., & Martin, R. P. (2007). Detecting and localizing wireless spoofing attacks. In *SECON'07* (pp. 193–202).
14. Chintalapudi, K., Padmanabha Iyer, A., & Padmanabhan, V. N. (2010). Indoor localization without the pain. In *Proceedings of the sixteenth annual international conference on Mobile computing and networking, MobiCom '10* (pp. 173–184). New York, NY: ACM.
15. Chrysikos, T., & Kotsopoulos, S. (2009). Impact of channel-dependent variation of path loss exponent on wireless information-theoretic security. In *Proceedings of the 2009 conference on Wireless telecommunications symposium, WTS'09* (pp. 384–390). Piscataway, NJ: IEEE Press. http://portal.acm.org/citation.cfm?id=1689059.1689128.
16. Conet roadmap. (2009). http://www.cooperating-objects.eu/roadmap/download/.
17. Costa, J. A., & Patwari, N. (2006). Distributed weighted-multidimensional scaling for node localization in sensor networks. *ACM Transaction on Sensor Networks*, *2*(1), 39–64. doi:10.1145/1138127.1138129.
18. Do, K. (2008). Effect of an additional sensor on aoa localization performance. *Simulation*, *11*(2), 1–5.
19. Drane, C., Macnaughtan, M., & Scott, C. (1998). Positioning gsm telephones. *IEEE Communications Magazine*, *36*(4), 46–54.
20. Efrat, A., Forrester, D., Iyer, A., Kobourov, S. G., Erten, C., & Kilic, O. (2010). Force-directed approaches to sensor localization. *ACM Transactions on Sensor Networks*, *7*, 271–2725. doi:10.1145/1807048.1807057.
21. Elnahraway, E., Li, X., & Martin, R. P. (2004). The limits of localization using rss. In *Proceedings of the 2nd international conference on embedded networked sensor systems, SenSys '04* (pp. 283–284). New York, NY: ACM. doi:10.1145/1031495.1031537.
22. Erol-Kantarci, M., Mouftah, H. T., & Oktug, S. (2011). A survey of architectures and localization techniques for underwater acoustic sensor networks. *IEEE Communications Surveys & Tutorials*, 1–16. doi:10.1109/SURV.2011.020211.00035.
23. Esposito, C., Cotroneo, D., & Ficco, M. (2009). Calibrating rss-based indoor positioning systems. In *Proceedings of the 2009 IEEE international conference on wireless and mobile computing, networking and communications, WIMOB '09* (pp. 1–6). Washington, DC: IEEE Computer Society. doi:10.1109/WiMob.2009.11.
24. Figueiras, J., & Frattasi, S. (2010). *Mobile positioning and tracking: From conventional to cooperative techniques*. London: Wiley. http://books.google.com/books?id=9J2siBo3LskC.
25. Goldoni, E., Savioli, A., Risi, M., & Gamba, P. (2010). *Experimental analysis of RSSI-based indoor localization with IEEE 802.15.4* (pp. 71–77). doi:10.1109/EW.2010.5483396.

26. Gotsman, C., & Koren, Y. (2004). Distributed graph layout for sensor networks. In J. Pach (Ed.), *Graph drawing, lecture notes in computer science* (Vol. 3383, pp. 273–284). Berlin: Springer.
27. Gu, Y., Lo, A., & Niemegeers, I. G. (2009). A survey of indoor positioning systems for wireless personal networks. *IEEE Communications Surveys & Tutorials*, *11*(1), 13–32.
28. Guvenc, I., & Chong, C. C. (2009). A survey on toa based wireless localization and nlos mitigation techniques. *IEEE Communications Surveys & Tutorials*, *11*, 107–124. doi:10.1109/SURV.2009.090308.
29. Harter, A., Hopper, A., Steggles, P., Ward, A., & Webster, P. (1999). The anatomy of a context-aware application. In *Proceedings of the 5th annual ACM/IEEE international conference on mobile computing and networking, MobiCom '99* (pp. 59–68). New York, NY: ACM. doi:10.1145/313451.313476.
30. He, T., Huang, C., John, B. M. B., & Stankovic Abdelzaher, T. (2003). Range-free localization schemes for large scale sensor, networks. doi:10.1145/938985.938995.
31. Ho, K., & Chan, Y. (1993). Solution and performance analysis of geolocation by tdoa. *IEEE Transactions on Aerospace and Electronic Systems*, *29*(4), 1311–1322. doi:10.1109/7.259534.
32. Honkavirta, V., Perala, T., Ali-Loytty, S., & Piche, R. (2009). A comparative survey of wlan location fingerprinting methods. In *The 6th workshop on positioning, navigation and. Communication* (pp. 243–251). doi:10.1109/WPNC.2009.4907834.
33. Jamâa, M. B., & Koubâa, A. (2012). Demo abstract: Rss-based localization in sensor networks does not need pre-deployment profiling. In *The 9th European conference on wireless sensor networks (EWSN)*.
34. Jasch, A., Feuerle, T., Scoor, G., & Hecker, P. (2010). *Geometrical siting considerations for wide area multilateration systems* (pp. 1304–1308). doi:10.1109/PLANS.2010.5507349.
35. Ji, X., & Zha, H. (2004). Sensor positioning in wireless ad-hoc sensor networks using multidimensional scaling. In *Proceedings of IEEE INFOCOM, Hong Kong, China* (pp. 2652–2661).
36. Jiao, L., Xing, J., & Li, F. Y. (2009). Performance comparison of residual related algorithms for toa positioning in wireless terrestrial and sensor networks. In *First international conference on wireless communication, vehicular technology, information theory and aerospace and electronic systems technology, 2009. Wireless VITAE* (pp. 278–283). doi:10.1109/WIRELESSVITAE.2009.5172462.
37. Kaemarungsi, K., & Krishnamurthy, P. (2004). Modeling of indoor positioning systems based on location fingerprinting. In *INFOCOM*.
38. Kaltiokallio, O., Bocca, M., & Patwari, N. (2012). Follow grandma: Long-term device-free localization for residential monitoring. In *Proceedings of the seventh IEEE international workshop on practical issues in building sensor networks applications, SenseApp 2012*.
39. Karl, H., & Willig, A. (2005). *Protocols and architectures for wireless sensor networks*. London: Wiley.
40. Kleinrock, L., & Silvester, J. (1978). Optimum transmission radii for packet radio networks or why six is a magic number. In *IEEE national telecommunications conference* (pp. 431–435).
41. Kushki, A., Plataniotis, K. N., & Venetsanopoulos, A. N. (2007). Kernel-based positioning in wireless local area networks. *IEEE Transcations on Mobile Computing*, *6*(6), 689–705.
42. Langendoen, K., & Reijers, N. (2003). Distributed localization in wireless sensor networks: A quantitative comparison. *Computer Networks, 43*, 499–518. doi:10.1016/S1389-1286(03)00356-6.
43. Lazos, L., & Poovendran, R. (2005). Serloc: Robust localization for wireless sensor networks. *ACM Transcations Sensor Networks*, *1*, 73–100. doi:10.1145/1077391.1077395.
44. Lederer, S., Wang, Y., & Gao, J. (2009). Connectivity-based localization of large-scale sensor networks with complex shape. *ACM Transcations Sensor Networks*, *5*(4), 1–32.
45. Li, L., & Kunz, T. (2009). Cooperative node localization using nonlinear data projection. *ACM Transcations Sensor Networks*, *5*(1), 1–26.
46. Li, M., & Liu, Y. (2007). Rendered path: Range-free localization in anisotropic sensor networks with holes. In *Proceedings of the 13th annual ACM international conference on mobile computing and networking, MobiCom '07* (pp. 51–62). New York, NY: ACM.
47. Li, X. (2009). Ratio-based zero-profiling indoor localization. In *IEEE 6th international conference on mobile adhoc and sensor systems, 2009. MASS '09* (pp. 40–49). doi:10.1109/MOBHOC.2009.5336964.
48. Li, X. (2009). Ratio-based zero-profiling indoor localization (pp. 40–49). In *MASS*.
49. Lim, H., Kung, L. C., Hou, J. C., & Luo, H. (2006). Zero-configuration, robust indoor localization: Theory and experimentation. In *INFOCOM*.
50. Lim, H., Kung, L. C., Hou, J. C., & Luo, H. (2010). Zero-configuration indoor localization over ieee 802.11 wireless infrastructure. *Wireless Networks*, *16*, 405–420. doi:10.1007/s11276-008-0140-3.
51. Liu, H., Darabi, H., Banerjee, P., & Liu, J. (2007). Survey of wireless indoor positioning techniques and systems. *IEEE Transaction on Systems, MAN, AND Cybernetics Part C Applications and Reviews*, *37*(6), 1067–1080.
52. Liu, N., Xu, Z., & Sadler, B. (2008). Low-complexity hyperbolic source localization with a linear sensor array. *IEEE Signal Processing Letters, 15*, 865–868.

53. Liu, X., Zhang, C., & Hu, J. (2008). Adaptive weights weighted centroid localization algorithm for wireless sensor networks. In *4th international conference on wireless communications, networking and mobile computing (WiCOM '08)* (pp. 1–4). doi:10.1109/WiCom.2008.851.
54. Mahalik, N. P. (2006). *Sensor networks and configuration: Fundamentals, standards, platforms, and applications*. Secaucus, NJ: Springer.
55. Mao, G., Anderson, B. D. O., & Fidan, B. (2007). Path loss exponent estimation for wireless sensor network localization. *Computer Networks*, *51*(10), 2467–2483. doi:10.1016/j.comnet.2006.11.007.
56. Mao, G., Fidan, B., & Anderson, B. D. O. (2007). Wireless sensor network localization techniques. *Computer Networks, 51*, 2529–2553. doi:10.1016/j.comnet.2006.11.018. http://dl.acm.org/citation.cfm?id=1242848.1243139.
57. Mao, G., Fidan, B., Mao, G., & Fidan, B. (2009). *Localization algorithms and strategies for wireless sensor networks*. Hershey, PA: Information Science Reference-Imprint of: IGI Publishing.
58. McCrady, D., Doyle, L., Forstrom, H., Dempsey, T., & Martorana, M. (2000). Mobile ranging using low-accuracy clocks. *IEEE Transactions on Microwave Theory and Techniques*, *48*(6), 951–958. doi:10.1109/22.846721.
59. Nagpal, R., Shrobe, H., & Bachrach, J. (2003). Organizing a global coordinate system from local information on an ad hoc sensor network. In *IPSN'03: Proceedings of the 2nd international conference on Information processing in sensor networks*.
60. Niculescu, D., & Badrinath, B. R. (2003). Ad hoc positioning system (aps) using aoa. In *INFOCOM*.
61. Official US government information about the global positioning system (gps) and related topics. http://www.gps.gov/.
62. Paul, A. S., & Wan, E. A. (2009). RSSI-based indoor localization and tracking using sigma-point kalman smoothers. *Journal of Selected Topics in Signal Processing*, *3*(5), 860–873.
63. Pi-Chun, C. (1999). A non-line-of-sight error mitigation algorithm in location estimation. In *Proceedings of the IEEE wireless communications and networking conference, WCNC* (pp. 59–68). doi:10.1109/WCNC.1999.797838.
64. Priyantha, N. B. (2005). *The cricket indoor location system*. Ph.D. thesis, Cambridge, MA, Adviser-Balakrishnan, Hari.
65. Priyantha, N. B., Balakrishnan, H., Demaine, E. D., & Teller, S. J. (2003). Anchor-free distributed localization in sensor networks. In *SenSys* (pp. 340–341).
66. Priyantha, N. B., Chakraborty, A., & Balakrishnan, H. (2000). The cricket location-support system. In *MobiCom '00: Proceedings of the 6th annual international conference on Mobile computing and networking* (pp. 32–43). New York, NY: ACM. doi:10.1145/345910.345917.
67. Rabbat, M., & Nowak, R. (2004). Distributed optimization in sensor networks. In *IPSN '04: Proceedings of the 3rd international symposium on information processing in sensor networks* (pp. 20–27). New York, NY: ACM. doi:10.1145/984622.984626.
68. Reichenbach, F., & Timmermann, D. (2006). Indoor localization with low complexity in wireless sensor networks. In *IEEE international conference on industrial informatics* (pp. 1018–1023). doi:10.1109/INDIN.2006.275737.
69. Ren, Y., Chuah, M. C., Yang, J., & Chen, Y. (2010). Detecting wormhole attacks in delay-tolerant networks. *Wireless Communications, 17*, 36–42. http://dl.acm.org/citation.cfm?id=1921927.1921936.
70. Römer, K. (2003). The lighthouse location system for smart dust. In *Proceedings of the 1st international conference on mobile systems, applications and services, MobiSys '03* (pp. 15–30). ACM: New York, NY. doi:10.1145/1066116.1189036.
71. Rong, P., & Sichitiu, M. (2006). Angle of arrival localization for wireless sensor networks. In *3rd Annual IEEE communications society on sensor and ad hoc communications and networks, 2006. SECON '06* (Vol. 1 pp. 374–382). doi:10.1109/SAHCN.2006.288442.
72. Rtrack project. http://www.rtrackp.com.
73. Shang, Y., Ruml, W., Zhang, Y., & Fromherz, M. P. J. (2003). Localization from mere connectivity. In *MobiHoc '03: Proceedings of the 4th ACM international symposium on mobile ad hoc networking & computing* (pp. 201–212). New York, NY: ACM. doi:10.1145/778415.778439.
74. Sources of errors in gps. (1999). http://www.kowoma.de/en/gps/errors.htm.
75. Stoleru, R., He, T., & Stankovic, J. A. (2004). Walking gps: A practical solution for localization in manually deployed wireless sensor networks. In *Annual IEEE conference on local computer networks* (pp. 480–489). doi:10.1109/LCN.2004.136.
76. Stoleru, R., He, T., & Stankovic, J. A. (2007). Range-free localization. In *Chapter in secure localization and time synchronization for wireless sensor and ad hoc networks*, Vol. 30.
77. Tateishi, K., & Ikegami, T. (2008). Estimation method of attenuation constant during localization in rssi. In *International symposium on communications and information technologies, ISCIT 2008* (pp. 482–487). doi:10.1109/ISCIT.2008.4700239.

78. Torrieri, D. (1984). Statistical theory of passive location systems. *IEEE Transactions on Aerospace and Electronic Systems*, *20*(2), 183–198. doi:10.1109/TAES.1984.310439.
79. Varshavsky, A., Pankratov, D., Krumm, J., & Lara, E. (2008). Calibree: Calibration-free localization using relative distance estimations. In *Proceedings of the 6th international conference on pervasive computing, pervasive '08* (pp. 146–161). Berlin: Springer.
80. Vidal, J., Najar, M., & Jativa, R. (2002). High resolution time-of-arrival detection for wireless positioning systems. In *Vehicular technology conference, 2002. Proceedings. VTC 2002-Fall. 2002 IEEE 56th*, Vol. 4 (pp. 2283–2287). doi:10.1109/VETECF.2002.1040627.
81. Wilson, J., & Patwari, N. (2010). Radio tomographic imaging with wireless networks. *IEEE Transactions on Mobile Computing*, *9*(5), 621–632.
82. Wymeersch, H., Lien, J., & Win, M. Z. (2009). Cooperative localization in wireless networks. In *Proceedings of the IEEE* (Vol. 97, pp. 427–450).
83. Yeh, L. W., Hsu, M. S., Lee, Y. F., & Tseng, Y. C. (2009). Indoor localization: Automatically constructing today's radio map by irobot and rfids. In *Sensors, 2009 IEEE* (pp. 1463–1466). doi:10.1109/ICSENS.2009.5398451.
84. Zhang, J., Yan, T., Stankovi, J. A., & Son, S. H. (2007). Thunder: Towards practical, zero cost acoustic localization for outdoor wireless sensor networks. *SIGMOBILE Mobile Computing and Communications Review*, *11*, 15–28. doi:10.1145/1234822.1234827.
85. Zhong, Z. (2010). Phd dissertation: Range-free localization and tracking in wireless sensor networks. http://www-users.cs.umn.edu/~zhong/papers/dissertation.pdf.
86. Zhu, J. (1992). Calculation of geometric dilution of precision. *IEEE Transactions on Aerospace and Electronic Systems*, *28*, 893–895.
87. Zhu, Y., Huang, D., & Jiang, A. (2008). Network localization using angle of arrival. In *IEEE international conference on electro/information technology* (pp. 205–210). doi:10.1109/EIT.2008.4554297.
88. Zorzi, F., & Zanella, A. (2009). Opportunistic localization: Modeling and analysis. In *VTC spring. IEEE*.